GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:09:37 ; Search time 26.082 Seconds
(without alignments)
1751.988 Million cell updates/sec

Title: US-10-683-610-1_COPY_1_104
Perfect score: 521
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....SDDEVSPREKQQTNSKGSSN 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	521	100.0	239	3	AAB56801	Aab56801 Human pro
2	521	100.0	530	7	ADB68448	Adb68448 Human DCA
3	521	100.0	530	7	ADB68450	Adb68450 Murine DC
4	521	100.0	530	7	ADB68452	Adb68452 Murine DC
5	521	100.0	530	7	ADB68446	Adb68446 Human DCA
6	521	100.0	530	8	ADM94432	Adm94432 Mouse ino
7	521	100.0	530	8	ABM80118	Abm80118 Tumour-as
8	521	100.0	530	8	ADP23045	Adp23045 PRO polyp
9	521	100.0	614	2	AAW56097	Aaw56097 Amino aci
10	367	70.4	500	2	AAW90061	Aaw90061 Human SAH
11	322	61.8	597	8	ADO43989	Ado43989 Amino aci
12	295	56.6	611	8	ABM80935	Abm80935 Tumour-as
13	288.5	55.4	455	4	ABG12029	Abg12029 Novel hum
14	273	52.4	564	5	ABG32371	Abg32371 Novel hum
15	268	51.4	508	8	ADS10638	Ads10638 Human the
16	268	51.4	509	8	ADL90193	Adl90193 Human enz
17	126.5	24.3	521	4	ABB58663	Abb58663 Drosophil
18	86.5	16.6	414	2	AAW10529	Aaw10529 Saccharom
19	86.5	16.6	414	4	AAB30809	Aab30809 Amino aci
20	86	16.5	419	8	ADR86387	Adr86387 Aspergill
21	86	16.5	446	8	ADR86528	Adr86528 Aspergill
22	86	16.5	584	8	ADR86515	Adr86515 Aspergill
23	85	16.3	1788	8	ADO17144	Ado17144 Novel exp
24	83.5	16.0	2058	8	ADM48570	Adm48570 Ankyrin r
25	83	15.9	628	8	ADS44146	Ads44146 Bacterial
26	82	15.7	2296	8	ADN60280	Adn60280 Human spl
27	81.5	15.6	428	5	ABG93245	Abg93245 C. albica
28	81	15.5	1499	4	AAM42018	Aam42018 Human pol
29	80	15.4	240	9	AEA28149	Aea28149 Sericin p
30	79.5	15.3	156	3	AAB43709	Aab43709 Human can
31	79.5	15.3	176	4	ABB71395	Abb71395 Drosophil
32	79.5	15.3	220	7	ADC33148	Adc33148 Human nov
33	79.5	15.3	1791	6	ABO52986	Abo52986 Human put
34	79.5	15.3	2752	8	ADN60278	Adn60278 Human ser
35	79.5	15.3	2752	8	ADP24102	Adp24102 PRO polyp
36	79.5	15.3	2752	9	ADY91619	Ady91619 Human pro
37	79	15.2	948	7	ADE34334	Ade34334 Retinobla
38	79	15.2	948	8	ADK67723	Adk67723 Human mod
39	79	15.2	948	8	ADK60419	Adk60419 Angiogene
40	79	15.2	948	8	ADK60720	Adk60720 Angiogene
41	79	15.2	948	8	ADP73343	Adp73343 Human ret
42	79	15.2	948	8	ADR51525	Adr51525 Human lip
43	79	15.2	948	9	ADX07074	Adx07074 Cyclin-de
44	79	15.2	973	4	AAM41045	Aam41045 Human pol
45	79	15.2	1278	4	AAM39259	Aam39259 Human pol

ALIGNMENTS

RESULT 1
AAB56801
ID AAB56801 standard; protein; 239 AA.
XX
AC AAB56801;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen protein sequence SEQ ID NO:1379.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.
XX
OS Homo sapiens.
XX
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005988.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587513/55.
DR N-PSDB; AAF16004.
XX
PT Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX
PS Claim 11; Page 1803-1804; 2338pp; English.
XX
CC AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen

CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 239 AA;

Query Match 100.0%; Score 521; DB 3; Length 239;
Best Local Similarity 100.0%; Pred. No. 5.1e-51;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 145

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db        146 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 189
```

RESULT 2
ADB68448
ID ADB68448 standard; protein; 530 AA.
XX
AC ADB68448;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human DCAL protein - SEQ ID 4.
XX
KW dendritic cell expressed S-adenosyl homocysteine hydrolase-like molecule;
KW DC; AHCY; DCAL; antiallergic; immunosuppressive; allergy; asthma;
KW allergic rhinitis; systemic anaphylaxis; autoimmune; diabetes mellitus;
KW rheumatoid arthritis; transplant rejection; vaccine; gene therapy; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1. .106
FT /note = Hydrophilic domain
FT Domain 107. .530
FT /note = adenosine-homocysteine hydrolase (AHCY)-like
FT domain
FT Binding-site 318. .323
FT /note= "NAD+ binding motif (GXGXXG)"
XX
PN WO2003055997-A1.
XX
PD 10-JUL-2003.
XX
PF 24-DEC-2002; 2002WO-AU001761.
XX
PR 24-DEC-2001; 2001AU-00009741.
XX
PA (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.
XX
PI Kato M, Angel NZ, Cooper BJ, Hart DNJ;
XX
DR WPI; 2003-559275/52.
DR N-PSDB; ADB68445, ADB68447.
XX
PT New dendritic cell expressed S-adenosyl homocysteine hydrolase-like
PT molecule (DCAL) gene, useful for preparing a composition for treating or
PT preventing a condition e.g., allergy, autoimmune disease or transplant
PT rejection.
XX
PS Disclosure; Fig 1B; 189pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising a
CC dendritic cell (DC)-expressed S-adenosyl homocysteine hydrolase (AHCY)-
CC like molecule (DCAL) gene. The polynucleotide of the invention
CC demonstrates antiallergic and immunosuppressive activities and may be
CC useful for treating a condition including an allergy such as asthma,
CC allergic rhinitis or systemic anaphylaxis, an autoimmune disease such as
CC diabetes mellitus or rheumatoid arthritis or transplant rejection.
CC Furthermore, the polynucleotide may be useful as a vaccine or during gene
CC therapy procedures. The genetically modified animal or the identified
CC modulatory agent of the invention may be used in the study of immunity,
CC DC function, brain physiology or neuronal cell function. The current
CC sequence is that of the human DCAL protein (SEQ ID 4) of the invention
CC which is encoded by the human DCAL cDNA.
XX
SQ Sequence 530 AA;

Query Match 100.0%; Score 521; DB 7; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.5e-50;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
```

RESULT 3
ADB68450
ID ADB68450 standard; protein; 530 AA.
XX
AC ADB68450;
XX
DT 04-DEC-2003 (first entry)
XX
DE Murine DCAL protein - SEQ ID 6.
XX
KW dendritic cell expressed S-adenosyl homocysteine hydrolase-like molecule;
KW DC; AHCY; DCAL; antiallergic; immunosuppressive; allergy; asthma;
KW allergic rhinitis; systemic anaphylaxis; autoimmune; diabetes mellitus;
KW rheumatoid arthritis; transplant rejection; vaccine; gene therapy;
KW murine; mouse.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Binding-site 318. .323
FT /note= "NAD+ binding motif (GXGXXG)"
XX
PN WO2003055997-A1.
XX
PD 10-JUL-2003.
XX
PF 24-DEC-2002; 2002WO-AU001761.
XX
PR 24-DEC-2001; 2001AU-00009741.
XX
PA (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.
XX
PI Kato M, Angel NZ, Cooper BJ, Hart DNJ;
XX
DR WPI; 2003-559275/52.
DR N-PSDB; ADB68449.
XX
PT New dendritic cell expressed S-adenosyl homocysteine hydrolase-like
PT molecule (DCAL) gene, useful for preparing a composition for treating or
PT preventing a condition e.g., allergy, autoimmune disease or transplant
PT rejection.
XX
PS Disclosure; Page 170-172; 189pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising a

```
CC   dendritic cell (DC)-expressed S-adenosyl homocysteine hydrolase (AHCY)-
CC   like molecule (DCAL) gene. The polynucleotide of the invention
CC   demonstrates antiallergic and immunosuppressive activities and may be
CC   useful for treating a condition including an allergy such as asthma,
CC   allergic rhinitis or systemic anaphylaxis, an autoimmune disease such as
CC   diabetes mellitus or rheumatoid arthritis or transplant rejection.
CC   Furthermore, the polynucleotide may be useful as a vaccine or during gene
CC   therapy procedures. The genetically modified animal or the identified
CC   modulatory agent of the invention may be used in the study of immunity,
CC   DC function, brain physiology or neuronal cell function. The current
CC   sequence is that of the murine DCAL protein (SEQ ID 6) of the invention
CC   which is encoded by genomic DNA.
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 521;  DB 7;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-50;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 4
ADB68452
ID   ADB68452 standard; protein; 530 AA.
XX
AC   ADB68452;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Murine DCAL protein - SEQ ID 8.
XX
KW   dendritic cell expressed S-adenosyl homocysteine hydrolase-like molecule;
KW   DC; AHCY; DCAL; antiallergic; immunosuppressive; allergy; asthma;
KW   allergic rhinitis; systemic anaphylaxis; autoimmune; diabetes mellitus;
KW   rheumatoid arthritis; transplant rejection; vaccine; gene therapy;
KW   murine; mouse.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Binding-site    318. .323
FT                   /note= "NAD+ binding motif (GXGXXG)"
XX
PN   WO2003055997-A1.
XX
PD   10-JUL-2003.
XX
PF   24-DEC-2002; 2002WO-AU001761.
XX
PR   24-DEC-2001; 2001AU-00009741.
XX
PA   (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.
XX
PI   Kato M,  Angel NZ,  Cooper BJ,  Hart DNJ;
XX
DR   WPI; 2003-559275/52.
DR   N-PSDB; ADB68451.
XX
PT   New dendritic cell expressed S-adenosyl homocysteine hydrolase-like
PT   molecule (DCAL) gene, useful for preparing a composition for treating or
PT   preventing a condition e.g., allergy, autoimmune disease or transplant
PT   rejection.
XX
PS   Disclosure; Page 175-177; 189pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising a
```

```
CC   dendritic cell (DC)-expressed S-adenosyl homocysteine hydrolase (AHCY)-
CC   like molecule (DCAL) gene. The polynucleotide of the invention
CC   demonstrates antiallergic and immunosuppressive activities and may be
CC   useful for treating a condition including an allergy such as asthma,
CC   allergic rhinitis or systemic anaphylaxis, an autoimmune disease such as
CC   diabetes mellitus or rheumatoid arthritis or transplant rejection.
CC   Furthermore, the polynucleotide may be useful as a vaccine or during gene
CC   therapy procedures. The genetically modified animal or the identified
CC   modulatory agent of the invention may be used in the study of immunity,
CC   DC function, brain physiology or neuronal cell function. The current
CC   sequence is that of the murine DCAL protein (SEQ ID 8) of the invention
CC   which is encoded by cDNA.
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 521;  DB 7;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-50;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 5
ADB68446
ID   ADB68446 standard; protein; 530 AA.
XX
AC   ADB68446;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human DCAL protein - SEQ ID 2.
XX
KW   dendritic cell expressed S-adenosyl homocysteine hydrolase-like molecule;
KW   DC; AHCY; DCAL; antiallergic; immunosuppressive; allergy; asthma;
KW   allergic rhinitis; systemic anaphylaxis; autoimmune; diabetes mellitus;
KW   rheumatoid arthritis; transplant rejection; vaccine; gene therapy; human;
KW   chromosome 1p21.1.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .106
FT                   /note = Hydrophilic domain
FT   Domain          107. .530
FT                   /note = adenosine-homocysteine hydrolase (AHCY)-like
FT                   domain
FT   Binding-site    318. .323
FT                   /note= "NAD+ binding motif (GXGXXG)"
XX
PN   WO2003055997-A1.
XX
PD   10-JUL-2003.
XX
PF   24-DEC-2002; 2002WO-AU001761.
XX
PR   24-DEC-2001; 2001AU-00009741.
XX
PA   (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.
XX
PI   Kato M,  Angel NZ,  Cooper BJ,  Hart DNJ;
XX
DR   WPI; 2003-559275/52.
DR   N-PSDB; ADB68445, ADB68447.
XX
PT   New dendritic cell expressed S-adenosyl homocysteine hydrolase-like
PT   molecule (DCAL) gene, useful for preparing a composition for treating or
PT   preventing a condition e.g., allergy, autoimmune disease or transplant
```

```
PT   rejection.
XX
PS   Disclosure; Fig 1B; 189pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising a
CC   dendritic cell (DC)-expressed S-adenosyl homocysteine hydrolase (AHCY)-
CC   like molecule (DCAL) gene. The polynucleotide of the invention
CC   demonstrates antiallergic and immunosuppressive activities and may be
CC   useful for treating a condition including an allergy such as asthma,
CC   allergic rhinitis or systemic anaphylaxis, an autoimmune disease such as
CC   diabetes mellitus or rheumatoid arthritis or transplant rejection.
CC   Furthermore, the polynucleotide may be useful as a vaccine or during gene
CC   therapy procedures. The genetically modified animal or the identified
CC   modulatory agent of the invention may be used in the study of immunity,
CC   DC function, brain physiology or neuronal cell function. The current
CC   sequence is that of the human DCAL protein (SEQ ID 2) of the invention
CC   which is encoded by DNA localised to chromosome 1p21.1.
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 521;  DB 7;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-50;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 6
ADM94432
ID   ADM94432 standard; protein; 530 AA.
XX
AC   ADM94432;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Mouse inositol 1,4,5-triphosphate receptor binding protein.
XX
KW   inositol-triphosphate receptor-binding protein; IP3 receptor;
KW   fluorescence resonance energy transfer; IP3 quantification;
KW   IP3-induced Ca2+ releasing activity; PIP2 function.
XX
OS   Mus musculus.
XX
PN   EP1408049-A2.
XX
PD   14-APR-2004.
XX
PF   13-OCT-2003; 2003EP-00256457.
XX
PR   11-OCT-2002; 2002JP-00299429.
XX
PA   (RIKE ) RIKEN KK.
XX
PI   Mikoshiba K,  Ando H,  Mizutani A,  Matsu-Ura T;
XX
DR   WPI; 2004-349314/33.
DR   N-PSDB; ADM94433.
XX
PT   New inositol 1,4,5-trisphosphate receptor-binding protein that binds to
PT   inositol 1,4,5-trisphosphate receptor, useful as inositol 1,4,5-
PT   trisphosphate indicator intracellularly and in cell-free system.
XX
PS   Claim 1; SEQ ID NO 1; 61pp; English.
XX
CC   The invention relates to a novel inositol 1,4,5-triphosphate (IP3)
CC   receptor-binding protein (I) that binds to IP3 receptor. (I) is used as
CC   an IP3 indicator in combination with a fluorescent molecule for detecting
CC   IP3 by the fluorescence resonance energy transfer (FRET) method using a
CC   combination of two fluorescent molecules applicable to the FRET method,
CC   and a protein containing at least IP3-binding domain of IP3 receptor
CC   labelled with the other fluorescent molecule. (I) is useful as IP3
CC   indicator intracellularly and in cell-free system, for detecting and/or
CC   quantifying IP3, as modulator of IP3-induced Ca2+ releasing activity,
CC   modulator of PIP2 function and Na2+ and bicarbonate ions. This sequence
CC   represents the IP3 receptor binding protein.
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 521;  DB 8;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-50;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 7
ABM80118
ID   ABM80118 standard; protein; 530 AA.
XX
AC   ABM80118;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Tumour-associated antigenic target (TAT) polypeptide PRO80619, SEQ:301.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   chromosome identification; chromosome mapping; gene mapping;
KW   gene therapy; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wu TD,  Zhang Z,  Zhou Y;
XX
DR   WPI; 2004-347921/32.
DR   N-PSDB; ACN37426.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
XX
PS   Claim 12; SEQ ID NO 301; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)
CC   polypeptides, and their related nucleic acids. The TAT polypeptides are
CC   overexpressed in cancer tissues compared to normal tissues, and may thus
CC   serve as effective targets for the diagnosis and treatment of cancer in
CC   mammals. The invention also relates to nucleic acid and polypeptide
CC   sequences at least 80% identical to the TAT nucleic acids and
CC   polypeptides; expression vectors and host cells comprising a TAT nucleic
```

```
CC   acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC   molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC   TAT polypeptide; and methods and compositions for the treatment or
CC   diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC   antibodies, antagonists, binding molecules and compositions are useful
CC   for diagnosing or treating a cell proliferative disorder associated with
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT polypeptide of the invention
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 521;  DB 8;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-50;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 8
ADP23045
ID   ADP23045 standard; protein; 530 AA.
XX
AC   ADP23045;
XX
DT   18-NOV-2004  (first entry)
XX
DE   PRO polypeptide SEQ ID NO:139.
XX
KW   PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW   osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW   antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX
OS   Unidentified.
XX
PN   WO2004041170-A2.
XX
PD   21-MAY-2004.
XX
PF   30-OCT-2003; 2003WO-US034312.
XX
PR   01-NOV-2002; 2002US-0423394P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Clark H,  Schoenfeld J,  Van Lookeren M,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-419628/39.
DR   N-PSDB; ADP23044.
XX
PT   New PRO polypeptides and polynucleotides, useful for treating e.g.
PT   erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT   renal disease, or demyelinating diseases of the central or peripheral
PT   nervous system.
XX
PS   Claim 7; SEQ ID NO 139; 2940pp; English.
XX
CC   The invention relates to a novel isolated nucleic acid and the PRO
CC   polypeptide encoded by it. A protein of the invention has
CC   antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC   osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC   antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC   of the invention may have a use in gene therapy. The PRO polypeptide, its
CC   agonist, antagonist, or antibody that specifically binds to the
CC   polypeptide is useful for treating an immune related disorder such as
CC   systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC   juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC   idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC   vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC   thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC   disease, a demyelinating disease of the central or peripheral nervous
CC   system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC   a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC   disease, infectious or autoimmune chronic active hepatitis, primary
CC   biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC   inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC   disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC   disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC   disease, asthma, allergic rhinitis, atopic dermatitis, food
CC   hypersensitivity, urticaria, an immunologic disease of the lung,
CC   eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC   pneumonitis, a transplantation associated disease, graft rejection or
CC   graft-versus-host disease. The present sequence represents a PRO protein
CC   of the invention.
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 521;  DB 8;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-50;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 9
AAW56097
ID   AAW56097 standard; protein; 614 AA.
XX
AC   AAW56097;
XX
DT   21-AUG-1998  (first entry)
XX
DE   Amino acid sequence of the 0DD4b5.3 enzyme.
XX
KW   DD4b9.3 enzyme; S-adenosyl-L-homocysteine hydrolase-type activity;
KW   AHCY-type activity; stimulation; inhibition; dendritic cell; cancer;
KW   autoimmune disease; transplantation.
XX
OS   Homo sapiens.
XX
PN   WO9814562-A1.
XX
PD   09-APR-1998.
XX
PF   06-OCT-1997;   97WO-NZ000133.
XX
PR   04-OCT-1996;   96NZ-00299507.
XX
PA   (HART/) HART D N J.
XX
PI   Hart DNJ;
XX
DR   WPI; 1998-240074/21.
DR   N-PSDB; AAV28617.
XX
PT   New isolated S-adenosyl-L-homocysteine hydrolase enzyme - is used to
PT   develop products which can be used in the treatment of e.g. auto-immune
PT   disease, transplantations or cancers.
XX
```

```
PS   Claim 3; Fig 1; 33pp; English.
XX
CC   This is the amino acid sequence of the DD4b9.3 enzyme which has S-
CC   adenosyl-L-homocysteine hydrolase (AHCY)-type activity. Its products can
CC   be used to identify substances which have a stimulatory or inhibitory
CC   effect on the enzyme activity. Such substances can be used to modulate
CC   dendritic cell (DC) function and for immunomodulation. They can be used
CC   in the treatment of e.g. autoimmune diseases, transplantations or
CC   cancers. The products can also be used for detection and diagnosis
XX
SQ   Sequence 614 AA;

  Query Match             100.0%;  Score 521;  DB 2;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-50;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 144

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db        145 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 188


RESULT 10
AAW90061
ID   AAW90061 standard; protein; 500 AA.
XX
AC   AAW90061;
XX
DT   04-MAR-1999  (first entry)
XX
DE   Human SAHH protein.
XX
KW   S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
KW   treatment; infection; cancer; autoimmune disease; detection; diagnosis;
KW   gene mapping; antisense; therapy; antagonist; immunoassay.
XX
OS   Homo sapiens.
XX
PN   US5854023-A.
XX
PD   29-DEC-1998.
XX
PF   17-JUL-1997;   97US-00896005.
XX
PR   17-JUL-1997;   97US-00896005.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Shah P,  Lal P,  Hillman JL,  Corley NC;
XX
DR   WPI; 1999-094906/08.
DR   N-PSDB; AAV73924.
XX
PT   Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase - for
PT   production of recombinant enzyme, useful for diagnosis, treatment and
PT   prevention of cancers, infections and autoimmune diseases.
XX
PS   Claim 1; Fig 1A-I; 40pp; English.
XX
CC   This sequence represents a human S-adenosyl-5-homocysteine hydrolase
CC   (SAHH). The SAHH protein can be used to generate specific antibodies and
CC   in drug screening to identify specific binding agents. Antagonists of the
CC   protein are used to treat or prevent a wide range of viral, bacterial,
CC   fungal, parasitic, protozoal or helminthic infections, many cancers
CC   (leukaemia, lymphoma or solid tumours), and many autoimmune diseases
CC   (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
CC   mellitus, multiple sclerosis etc). All these conditions may be treated by
CC   expressing antisense sequences, triplex-forming agents or ribozymes
CC   directed against the nucleic acid. The nucleic acid and its fragments can
CC   be used as probes or primers for detecting and quantifying gene
CC   expression, for diagnosis or monitoring of disease, to identify genetic
CC   variations, mutations or polymorphisms, in gene mapping and as antisense
CC   therapeutics. Antibodies are used directly as antagonists, indirectly to
CC   deliver active agents to SAHH-expressing cells, to diagnose and monitor
CC   diseases in standard immunoassays, in competitive drug screens and to
CC   isolate the protein from natural sources
XX
SQ   Sequence 500 AA;

  Query Match             70.4%;  Score 367;  DB 2;  Length 500;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-33;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         31 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 60

Qy         91 PREKQQTNSKGSSN 104
              ||||||||||||||
Db         61 PREKQQTNSKGSSN 74


RESULT 11
ADO43989
ID   ADO43989 standard; protein; 597 AA.
XX
AC   ADO43989;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Amino acid sequence of human putative S-adenosylhomocysteinase.
XX
KW   protein complex; neurological disease; stroke; neurodegeneration;
KW   Wallerian degeneration; Alzheimer's disease; neurological disorder;
KW   epilepsy; inflammatory condition; ulcerative colitis; Crohn's disease;
KW   atherosclerosis; 1D-MYO-inositol triphosphate 3 kinase A; ASK1; ASK2;
KW   ASK3; CaMKII beta; CaMKII delta; CaMKII gamma; casein kinase II alpha;
KW   Cdc37; CHK2; CTCL tumour antigen SE20-4; EF-1 alpha 1; EMAP;
KW   FLJ14653 NT2RP2002252; FLJ30839 FEBRA2002429; HERC2;
KW   inositol polyphosphate-5-phosphatase; inositol-1; 4;
KW   5-triphosphate 5-phosphatase type 1; IRAK1; IRAK4; KIAA1441; MSTP030;
KW   Nek9; PAR3; Pellino 1; Pellino 3; podocalyxin-like protein 1 precursor;
KW   Pushover; S-adenosylhomocysteinase;
KW   secretory carrier-associated membrane protein 2; surfeit locus protein 2;
KW   ubiquitin carboxyl terminal hydrolase 11;
KW   upstream regulatory element binding protein 1; Vartul;
KW   Werner's syndrome helicase interacting protein; WHIP;
KW   X-ray repair cross complementing protein 4.
XX
OS   Homo sapiens.
XX
PN   WO2004031242-A2.
XX
PD   15-APR-2004.
XX
PF   11-SEP-2003; 2003WO-EP010110.
XX
PR   12-SEP-2002; 2002EP-00020495.
PR   12-SEP-2002; 2002EP-00020496.
PR   12-SEP-2002; 2002EP-00020497.
XX
PA   (CELL-) CELLZOME AG.
XX
PI   Bouwmeester T,  Drewes G,  Jackson D,  Helftenbein G,  Schirle M;
PI   Kuester B,  Hopf C;
XX
DR   WPI; 2004-316467/29.
XX
PT   New complex comprising at least one first protein, and at least one
PT   second protein, useful for treating stroke, Alzheimer's disease,
PT   neurological disorders such as epilepsy, and inflammatory conditions such
PT   as ulcerative colitis.
XX
```

```
PS   Example; Page 191-193; 287pp; English.
XX
CC   The specification describes protein complexes involved in cellular
CC   processes which have been shown to be critical for the development of
CC   various forms of neurological diseases. Three protein complexes were
CC   identified: ASK2 protein complex, Pellino-1 protein complex and Pellino-3
CC   protein complex. The protein complex are useful for treating diseases and
CC   disorders, e.g. stroke, neurodegeneration such as Wallerian degeneration,
CC   Alzheimer's disease, neurological disorders such as epilepsy, and
CC   inflammatory conditions such as ulcerative colitis, Crohn's disease or
CC   atherosclerosis. Proteins identified as being part of the protein
CC   complexes of the invention are 1D-MYO-inositol triphosphate 3 kinase A,
CC   ASK1, ASK2, ASK3, CaMKII beta, CaMKII delta, CaMKII gamma, casein kinase
CC   II alpha, Cdc37, CHK2, CTCL tumour antigen SE20-4, EF-1 alpha 1, EMAP,
CC   FLJ14653 NT2RP2002252, FLJ30839 FEBRA2002429, HERC2, two hypothetical
CC   proteins of 35.5 Kda and 49.3 Kda, inositol polyphosphate-5-phosphatase,
CC   inositol-1,4,5-triphosphate 5-phosphatase type 1, IRAK1, IRAK4, KIAA1441,
CC   MSTP030, Nek9, PAR3, Pellino 1, Pellino 3, podocalyxin-like protein 1
CC   precursor, Pushover, a putative S-adenosylhomocysteinase, secretory
CC   carrier-associated membrane protein 2, surfeit locus protein 2, ubiquitin
CC   carboxyl terminal hydrolase 11, upstream regulatory element binding
CC   protein 1, Vartul, Werner's syndrome helicase interacting protein (WHIP),
CC   X-ray repair cross complementing protein 4 (isoform 1). The present
CC   sequence represents putative S-adenosylhomocysteinase.
XX
SQ   Sequence 597 AA;

  Query Match             61.8%;  Score 322;  DB 8;  Length 597;
  Best Local Similarity   54.7%;  Pred. No. 1.1e-27;
  Matches   76;  Conservative    8;  Mismatches   19;  Indels   36;  Gaps    4;

Qy          1 MSMPDAMPLP-------GVGE---ELKQAKEIEDAEKYSFMATVTKAPKK---------- 40
              :| | |:|||       |  |    |:| :  | ::    :  :  |
Db         34 ISQP-ALPLPQETQAPSGESENLRNLRQGRRAESSQNLLSLRRLVSAQGNGRSDWRWLWI 92

Qy         41 ---------------QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSS 85
                            ||||||||||||||||||||||||||||||||||||||||||||
Db         93 QLWPEIDSLEISCFLQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSS 152

Qy         86 DDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||
Db        153 DDEVSPREKQQTNSKGSSN 171


RESULT 12
ABM80935
ID   ABM80935 standard; protein; 611 AA.
XX
AC   ABM80935;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Tumour-associated antigenic target (TAT) polypeptide PRO81608, SEQ:2410.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   chromosome identification; chromosome mapping; gene mapping;
KW   gene therapy; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wu TD,  Zhang Z,  Zhou Y;
XX
DR   WPI; 2004-347921/32.
DR   N-PSDB; ACN38718.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
XX
PS   Claim 12; SEQ ID NO 2410; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)
CC   polypeptides, and their related nucleic acids. The TAT polypeptides are
CC   overexpressed in cancer tissues compared to normal tissues, and may thus
CC   serve as effective targets for the diagnosis and treatment of cancer in
CC   mammals. The invention also relates to nucleic acid and polypeptide
CC   sequences at least 80% identical to the TAT nucleic acids and
CC   polypeptides; expression vectors and host cells comprising a TAT nucleic
CC   acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC   molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC   TAT polypeptide; and methods and compositions for the treatment or
CC   diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC   antibodies, antagonists, binding molecules and compositions are useful
CC   for diagnosing or treating a cell proliferative disorder associated with
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT polypeptide of the invention
XX
SQ   Sequence 611 AA;

  Query Match             56.6%;  Score 295;  DB 8;  Length 611;
  Best Local Similarity   55.2%;  Pred. No. 1.4e-24;
  Matches   69;  Conservative   11;  Mismatches   17;  Indels   28;  Gaps    3;

Qy          8 PLPGVGE-----------ELKQAKEIEDAEKY--------------SFMATVTKAP---K 39
              |:|| |             :: ||  :: :: :              |   |||:||   |
Db         61 PVPGPGSGPAAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 120

Qy         40 KQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNS 99
              |||||||  ||| | ||| ||||||||||||||||||||||||||||| |||:||| ||
Db        121 KQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNS 180

Qy        100 KGSSN 104
              |||||:
Db        181 KGSSD 185


RESULT 13
ABG12029
ID   ABG12029 standard; protein; 455 AA.
XX
AC   ABG12029;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #12020.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
```

```
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS76216.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 42388; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC   amino acid sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 455 AA;

  Query Match             55.4%;  Score 288.5;  DB 4;  Length 455;
  Best Local Similarity   54.0%;  Pred. No. 5.3e-24;
  Matches   67;  Conservative   13;  Mismatches   17;  Indels   27;  Gaps    3;

Qy          8 PLPGVGE----------ELKQAKEIEDAEKY-------------SFMATVTKAPK--K 40
              |:|| |            :: ||  :: :: :             |   |||:||:  |
Db         61 PVPGPGSGPAAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 120

Qy         41 QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSK 100
              :|||||  ||| | ||| ||||||||||||||||||||||||||||| |||:||| |||
Db        121 KIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSK 180

Qy        101 GSSN 104
              |||:
Db        181 GSSD 184


RESULT 14
ABG32371
ID   ABG32371 standard; protein; 564 AA.
XX
AC   ABG32371;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Novel human enzyme, NZMS-1, Incyte 7484737CD1.
XX
KW   Human; NZMS-1; immune system disorder; AIDS; contact dermatitis;
KW   acquired immune deficiency syndrome; Chediak-Higashi syndrome;
KW   common variable immunodeficiency; developmental disorder; cataract;
KW   Cushing's syndrome; hypothyroidism; eye disorder; glaucoma; asthma;
KW   metabolic disorder; cystic fibrosis; sickle-cell anaemia; epilepsy;
KW   smooth muscle disorder; lactic acidosis; neurological disorder;
KW   Parkinson's disease; cardiac disorder; congestive heart failure;
KW   myocardial infarction; pulmonary disorder; emphysema; bronchitis;
KW   parasitic infection; amoebiasis; filariasis; cancer; leukaemia;
KW   cell proliferative disorder; Incyte 7484737CD1; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO200264795-A2.
XX
PD   22-AUG-2002.
XX
PF   08-FEB-2002; 2002WO-US003814.
XX
PR   09-FEB-2001; 2001US-0268113P.
PR   15-FEB-2001; 2001US-0269215P.
PR   27-FEB-2001; 2001US-0272271P.
PR   07-MAR-2001; 2001US-0274091P.
PR   09-MAR-2001; 2001US-0274423P.
PR   23-MAR-2001; 2001US-0278479P.
PR   23-MAR-2001; 2001US-0278480P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sanjanwala MM,  Lu Y,  Lee EA,  Hafalia AJA,  Warren BA,  Baughn MR;
PI   Tang TY,  Yue H,  Yao MG,  Lee S,  Thornton M,  Walia NK,  Xu Y;
PI   Tran UK,  Lal PG,  Lu DAM,  Swarnaker A,  Jones KA,  Ring HZ;
XX
DR   WPI; 2002-643488/69.
DR   N-PSDB; ABK91312.
XX
PT   New isolated human enzymes (NZMS), useful for diagnosing, treating and
PT   preventing immune system, developmental, eye, metabolic, smooth muscle,
PT   neurological, pulmonary and cardiac disorders, cancer, or parasitic
PT   infections.
XX
PS   Claim 1; Page 147-149; 170pp; English.
XX
CC   The invention relates to a new isolated human enzyme (NZMS) (I) and the
CC   encoding nucleic acid (II). The polypeptides and polynucleotides are
CC   useful in assessing the effects of exogenous compounds on the expression
CC   of nucleic acid and amino acid sequences of NZMS. The NZMS or its
CC   fragments are useful in screening compounds for effectiveness as agonist
CC   or antagonist of the polypeptides, or in altering the expression of the
CC   target polynucleotide and compounds that specifically bind to or modulate
CC   the activity of the polypeptide. The antibody is useful in diagnosing a
CC   condition or disease associated with the decreased expression or
CC   overexpression of NZMS and in detecting (I). (I) is useful in the
CC   production of antibodies. The microarray is useful in monitoring or
CC   measuring gene expression profiles. The polypeptides and polynucleotides
CC   are useful in diagnosing, treating and preventing diseases or conditions
CC   associated with the decreased expression or overexpression of NZMS, such
CC   as immune system disorders (e.g. acquired immune deficiency syndrome
CC   (AIDS), contact dermatitis), immune deficiencies (e.g. common variable
CC   immunodeficiency, Chediak-Higashi syndrome), developmental disorders
CC   (e.g. Cushing's syndrome, hypothyroidism), eye disorders (e.g. glaucoma,
CC   cataract), metabolic disorders (e.g. cystic fibrosis, sickle-cell
CC   anaemia), smooth muscle disorders (e.g. asthma, lactic acidosis),
CC   neurological disorders (e.g. Parkinson's disease, epilepsy), cardiac
CC   disorders (e.g. congestive heart failure, myocardial infarction),
CC   pulmonary disorders (emphysema, bronchitis), parasitic infections
CC   (amoebiasis, filariasis), and cell proliferative disorders (e.g. cancer,
CC   leukaemia). The microarray is useful in monitoring or measuring protein-
CC   protein interactions, drug-target interactions, and gene expression
CC   profiles. The present sequence represents the amino acid sequence of one
CC   of 11 novel human enzymes (NZMS1-11) of the invention
XX
SQ   Sequence 564 AA;

  Query Match             52.4%;  Score 273;  DB 5;  Length 564;
  Best Local Similarity   73.1%;  Pred. No. 4.2e-22;
  Matches   57;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;
```

```
Qy          27 KYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD 86
               |  ::       | ||||||  ||| | ||| ||||||||||||||||||||||||||
Db          61 KKKYIVNGNSGIKAQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSD 120

Qy          87 DEVSPREKQQTNSKGSSN 104
               || |||:||| ||||||:
Db         121 DETSPRDKQQKNSKGSSD 138


RESULT 15
ADS10638
ID   ADS10638 standard; protein; 508 AA.
XX
AC   ADS10638;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Human therapeutic protein - SEQ ID 875.
XX
KW   antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
KW   inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
KW   aplastic anaemia; cancer; wound healing; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2004080148-A2.
XX
PD   23-SEP-2004.
XX
PF   30-SEP-2003; 2003WO-US030720.
XX
PR   02-OCT-2002; 2002US-0416186P.
XX
PA   (NUVE-) NUVELO INC.
XX
PI   Tang YT,  Asundi V,  Ren F,  Zhang J,  Wehrman T,  Wang Z,  Ma Y;
PI   Wang D,  Chen R,  Zhao QA,  Wang J,  Ghosh M,  Xue AJ,  Weng G,  Zhou P;
XX
DR   WPI; 2004-668857/65.
DR   N-PSDB; ADS09954.
XX
PT   New polynucleotide, useful in preparing a composition for diagnosing or
PT   treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT   aplastic anemia or cancer for promoting wound healing.
XX
PS   Claim 20; SEQ ID NO 875; 718pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide and the encoded
CC   polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC   neuroprotective, antianaemic, cytostatic and vulnerary activities and may
CC   be useful in preparing a composition for diagnosing or treating
CC   inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC   disorders, such as aplastic anaemia or cancer, as well as for promoting
CC   wound healing. The molecules may also be utilised during gene therapy
CC   procedures. The current sequence is that of a human therapeutic protein
CC   of the invention. The current sequence is not shown explicitly within the
CC   specification but can be accessed from the WIPO web-site.
XX
SQ   Sequence 508 AA;

  Query Match             51.4%;  Score 268;  DB 8;  Length 508;
  Best Local Similarity   82.1%;  Pred. No. 1.4e-21;
  Matches   55;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          38 PKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQT 97
               |:  |||||  ||| | ||| |||||||||||||||||||||||||||||| |||:|||
Db          16 PEWMIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQK 75

Qy          98 NSKGSSN 104
               |||||||:
Db          76 NSKGSSD 82
```

Search completed: December 14, 2005, 08:16:42
Job time : 28.082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:20:23 ; Search time 121.215 Seconds
(without alignments)
1826.923 Million cell updates/sec

Title: US-10-683-610-1
Perfect score: 2781
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....QAKYLGLNKNGPFKPNYYRY 530

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2781	100.0	530	5	US-10-876-285-2	Sequence 2, Appli
2	2781	100.0	530	5	US-10-876-285-4	Sequence 4, Appli
3	2781	100.0	530	5	US-10-876-285-6	Sequence 6, Appli
4	2781	100.0	530	5	US-10-683-610-1	Sequence 1, Appli
5	2781	100.0	614	3	US-09-782-051-2	Sequence 2, Appli
6	2778	99.9	530	5	US-10-876-285-8	Sequence 8, Appli
7	2449.5	88.1	619	5	US-10-756-149-5569	Sequence 5569, Ap
8	2429	87.3	564	5	US-10-467-903-1	Sequence 1, Appli
9	1970	70.8	521	6	US-11-097-143-2781	Sequence 2781, Ap
10	1819.5	65.4	504	6	US-11-097-143-41922	Sequence 41922, A
11	1601.5	57.6	455	5	US-10-450-763-42388	Sequence 42388, A
12	1289	46.4	426	4	US-10-369-493-10135	Sequence 10135, A
13	1187	42.7	461	4	US-10-369-493-4783	Sequence 4783, Ap
14	1187	42.7	461	4	US-10-369-493-7543	Sequence 7543, Ap
15	1185	42.6	463	4	US-10-369-493-19704	Sequence 19704, A
16	1158	41.6	466	4	US-10-369-493-12110	Sequence 12110, A
17	1154.5	41.5	432	6	US-11-097-143-39477	Sequence 39477, A
18	1151.5	41.4	457	3	US-09-925-301-1059	Sequence 1059, Ap
19	1145.5	41.2	433	4	US-10-433-802-4	Sequence 4, Appli
20	1142	41.1	461	4	US-10-369-493-11466	Sequence 11466, A
21	1142	41.1	461	4	US-10-369-493-14652	Sequence 14652, A
22	1142	41.1	461	4	US-10-369-493-21032	Sequence 21032, A
23	1142	41.1	464	4	US-10-369-493-15125	Sequence 15125, A
24	1141	41.0	437	4	US-10-369-493-5097	Sequence 5097, Ap
25	1132	40.7	485	4	US-10-767-701-47244	Sequence 47244, A
26	1129	40.6	433	4	US-10-369-493-2425	Sequence 2425, Ap
27	1129	40.6	433	4	US-10-369-493-2430	Sequence 2430, Ap
28	1127	40.5	466	4	US-10-369-493-8750	Sequence 8750, Ap
29	1126	40.5	447	4	US-10-369-493-3683	Sequence 3683, Ap
30	1126	40.5	469	4	US-10-314-657-71	Sequence 71, Appl
31	1126	40.5	469	5	US-10-473-193-71	Sequence 71, Appl
32	1125	40.5	463	4	US-10-369-493-20829	Sequence 20829, A
33	1118	40.2	428	4	US-10-369-493-10445	Sequence 10445, A
34	1117	40.2	465	4	US-10-369-493-17934	Sequence 17934, A
35	1111	39.9	485	4	US-10-156-761-12587	Sequence 12587, A
36	1110	39.9	476	4	US-10-369-493-8266	Sequence 8266, Ap
37	1106	39.8	463	4	US-10-369-493-16776	Sequence 16776, A
38	1105	39.7	449	4	US-10-369-493-1740	Sequence 1740, Ap
39	1105	39.7	463	4	US-10-369-493-7922	Sequence 7922, Ap
40	1099.5	39.5	466	4	US-10-369-493-19391	Sequence 19391, A
41	1097	39.4	459	4	US-10-369-493-10671	Sequence 10671, A
42	1085	39.0	468	4	US-10-369-493-9241	Sequence 9241, Ap
43	1082	38.9	468	4	US-10-369-493-9386	Sequence 9386, Ap
44	1074	38.6	474	3	US-09-738-626-4336	Sequence 4336, Ap
45	1074	38.6	478	3	US-09-746-660A-98	Sequence 98, Appl

ALIGNMENTS

RESULT 1
US-10-876-285-2
; Sequence 2, Application US/10876285
; Publication No. US20050044582A1
; GENERAL INFORMATION:
; APPLICANT: Kato, Masato
; APPLICANT: Angel, Nicola Zoe
; APPLICANT: Cooper, Benjamine Jay
; APPLICANT: Hart, Derek Nigel John
; TITLE OF INVENTION: Novel Polynucleotides and uses therefor
; FILE REFERENCE: 10338-18US
; CURRENT APPLICATION NUMBER: US/10/876,285
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: PR9741/01
; PRIOR FILING DATE: 2001-12-24
; PRIOR APPLICATION NUMBER: PCT/AU02/01761
; PRIOR FILING DATE: 2002-12-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 530
; TYPE: PRT
; ORGANISM: human
US-10-876-285-2

Query Match 100.0%; Score 2781; DB 5; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.4e-244;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 2
US-10-876-285-4
; Sequence 4, Application US/10876285
; Publication No. US20050044582A1
; GENERAL INFORMATION:
;  APPLICANT: Kato, Masato
;  APPLICANT:  Angel, Nicola Zoe
;  APPLICANT:  Cooper, Benjamine Jay
;  APPLICANT:  Hart, Derek Nigel John
;  TITLE OF INVENTION: Novel Polynucleotides and uses therefor
;  FILE REFERENCE: 10338-18US
;  CURRENT APPLICATION NUMBER: US/10/876,285
;  CURRENT FILING DATE:  2004-06-24
;  PRIOR APPLICATION NUMBER: PR9741/01
;  PRIOR FILING DATE: 2001-12-24
;  PRIOR APPLICATION NUMBER: PCT/AU02/01761
;  PRIOR FILING DATE: 2002-12-24
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: Human
US-10-876-285-4

  Query Match           100.0%;  Score 2781;  DB 5;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-244;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 3
US-10-876-285-6
; Sequence 6, Application US/10876285
; Publication No. US20050044582A1
; GENERAL INFORMATION:
;  APPLICANT: Kato, Masato
;  APPLICANT:  Angel, Nicola Zoe
;  APPLICANT:  Cooper, Benjamine Jay
;  APPLICANT:  Hart, Derek Nigel John
;  TITLE OF INVENTION: Novel Polynucleotides and uses therefor
;  FILE REFERENCE: 10338-18US
;  CURRENT APPLICATION NUMBER: US/10/876,285
;  CURRENT FILING DATE:  2004-06-24
;  PRIOR APPLICATION NUMBER: PR9741/01
;  PRIOR FILING DATE: 2001-12-24
;  PRIOR APPLICATION NUMBER: PCT/AU02/01761
;  PRIOR FILING DATE: 2002-12-24
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: Mouse
US-10-876-285-6

  Query Match           100.0%;  Score 2781;  DB 5;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-244;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
```

```
RESULT 4
US-10-683-610-1
; Sequence 1, Application US/10683610
; Publication No. US20050181450A1
; GENERAL INFORMATION:
;  APPLICANT: RIKEN
;  TITLE OF INVENTION: A novel inositol 1,4,5-trisphosphate (IP3) receptor binding prote
;  TITLE OF INVENTION:  IP3
;  TITLE OF INVENTION:  indicator
;  FILE REFERENCE: PH-1852
;  CURRENT APPLICATION NUMBER: US/10/683,610
;  CURRENT FILING DATE:  2003-10-10
;  PRIOR APPLICATION NUMBER: JP 2002-299429
;  PRIOR FILING DATE: 2002-10-11
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-683-610-1

  Query Match            100.0%;  Score 2781;  DB 5;  Length 530;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-244;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 5
US-09-782-051-2
; Sequence 2, Application US/09782051
; Patent No. US20020035078A1
; GENERAL INFORMATION:
;  APPLICANT: Hart, Derek N J
;  TITLE OF INVENTION: Enzyme having S-adenosyl-L-homocysteine hydrolase
;  TITLE OF INVENTION:  (AHCY) type activity
;  FILE REFERENCE: 24305 MRB
;  CURRENT APPLICATION NUMBER: US/09/782,051
;  CURRENT FILING DATE:  2001-02-14
;  PRIOR APPLICATION NUMBER: PCT/NZ97/00133
;  PRIOR FILING DATE: 1997-10-06
;  PRIOR APPLICATION NUMBER: NZ 299507
;  PRIOR FILING DATE: 1996-10-04
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 614
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-782-051-2

  Query Match            100.0%;  Score 2781;  DB 3;  Length 614;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-244;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 144

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 204

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 264

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 324

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 384

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 444

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 504

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 564

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 614


RESULT 6
US-10-876-285-8
; Sequence 8, Application US/10876285
; Publication No. US20050044582A1
; GENERAL INFORMATION:
;  APPLICANT: Kato, Masato
;  APPLICANT:  Angel, Nicola Zoe
;  APPLICANT:  Cooper, Benjamine Jay
;  APPLICANT:  Hart, Derek Nigel John
;  TITLE OF INVENTION: Novel Polynucleotides and uses therefor
;  FILE REFERENCE: 10338-18US
;  CURRENT APPLICATION NUMBER: US/10/876,285
;  CURRENT FILING DATE:  2004-06-24
;  PRIOR APPLICATION NUMBER: PR9741/01
;  PRIOR FILING DATE: 2001-12-24
;  PRIOR APPLICATION NUMBER: PCT/AU02/01761
;  PRIOR FILING DATE: 2002-12-24
```

```
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: Mouse
US-10-876-285-8

  Query Match             99.9%;  Score 2778;  DB 5;  Length 530;
  Best Local Similarity   99.8%;  Pred. No. 2.6e-244;
  Matches  529;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||:|||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPTFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 7
US-10-756-149-5569
; Sequence 5569, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
;  TITLE OF INVENTION:  METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
;  FILE REFERENCE: file
;  CURRENT APPLICATION NUMBER: US/10/756,149
;  CURRENT FILING DATE:  2004-01-12
;  NUMBER OF SEQ ID NOS: 5818
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5569
;   LENGTH: 619
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-756-149-5569

  Query Match             88.1%;  Score 2449.5;  DB 5;  Length 619;
  Best Local Similarity   84.7%;  Pred. No. 3.3e-214;
  Matches  466;  Conservative   33;  Mismatches   24;  Indels   27;  Gaps    3;

Qy          8 PLPGVGE----------ELKQAKEIEDAEKY--------------SFMATVTKAPK--K 40
              |:|| |            :: ||  :: :: :              |   |||:||:  |
Db         70 PVPGPGSGPAAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 129

Qy         41 QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSK 100
              :|||||  ||| | ||| ||||||||||||||||||||||||||||| |||:||| |||
Db        130 KIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSK 189

Qy        101 GSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAV 160
              |||:|||||||||||||||||||:| ||::|||||||||||||||||||||||||||||
Db        190 GSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGAKIVGCTHITAQTAV 249

Qy        161 LIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDG 220
              |:||| |||||||:|||||| |||||||||:|  |||||||||||||||||||::|
Db        250 LMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGESEDDFWWCIDRCVNVEG 309

Qy        221 WQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVND 280
              || ||||||||||||:|||||:||||:|||||||||||||||||||||||||||||||||
Db        310 WQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLSKAGKLCVPAMNVND 369

Qy        281 SVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYIT 340
              |||||||||||||||||||||||:|||||||||||||||||||||||||||:|:|||:|
Db        370 SVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGCCAALKAMGSIVYVT 429

Qy        341 EIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHS 400
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        430 EIDPICALQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLDRMKNSCIVCNMGHS 489

Qy        401 NTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITAT 460
              |||||| ||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        490 NTEIDVASLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLNLSCSTVPTFVLSITAT 549

Qy        461 TQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKN 520
              ||||||||||||||||||||||||||||||||||||:|||||||||:|||||||||||
Db        550 TQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTELTDEQAKYLGLNKN 609

Qy        521 GPFKPNYYRY 530
              ||||||||||
Db        610 GPFKPNYYRY 619


RESULT 8
US-10-467-903-1
; Sequence 1, Application US/10467903
; Publication No. US20040265807A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE CORPORATION
;  APPLICANT:  SANJANWALA, Madhusudan M.
;  APPLICANT:  LU, Yan
;  APPLICANT:  LEE, Ernestine A.
;  APPLICANT:  HAFALIA, April J.A.
;  APPLICANT:  WARREN, Bridget A.
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  YUE, Henry
;  APPLICANT:  YAO, Monique G.
;  APPLICANT:  LEE, Sally
;  APPLICANT:  THORNTON, Michael B.
;  APPLICANT:  CHAWLA, Narinder K.
;  APPLICANT:  XU, Yuming
;  APPLICANT:  TRAN, Uyen K.
;  APPLICANT:  LAL, Preeti G.
;  APPLICANT:  LU, Dyung Aina M.
;  APPLICANT:  SWARNAKAR, Anita
;  APPLICANT:  RING, Huijun Z.
;  APPLICANT:  JONES, Karen A.
;  TITLE OF INVENTION: ENZYMES
;  FILE REFERENCE: PI-0360 USN
;  CURRENT APPLICATION NUMBER: US/10/467,903
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: PCT/US02/03814
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/268,113
```

; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/269,215
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US 60/272,271
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 60/274,091
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/274,423
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/278,480
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/278,479
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7484737CD1
US-10-467-903-1

Query Match 87.3%; Score 2429; DB 5; Length 564;
Best Local Similarity 90.3%; Pred. No. 2.1e-212;
Matches 455; Conservative 24; Mismatches 25; Indels 0; Gaps 0;

```
Qy          27 KYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD 86
               |  ::       | ||||||  ||| | ||| ||||||||||||||||||||||||||||
Db          61 KKKYIVNGNSGIKAQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSD 120

Qy          87 DEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGA 146
               || |||:||| |||||||:|||||||||||||||||||:| ||::|||||||||||||||
Db         121 DETSPRDKQQKNSKGSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGA 180

Qy         147 KIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESED 206
               |||||||||||||||:||| |||||||||:||||||| |||||||||:|  ||||||||||
Db         181 KIVGCTHITAQTAVLMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGESED 240

Qy         207 DFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLS 266
               ||||||||||::||| |||||||||||||:||||||:||||:||||||||||||||||||
Db         241 DFWWCIDRCVNVEGWQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLS 300

Qy         267 KAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGC 326
               ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db         301 KAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGC 360

Qy         327 CAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLD 386
               ||||||:|:|||:|||||||||||||||:||||||||:||||||||||||||||||||||
Db         361 CAALKAMGSIVYVTEIDPICALQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLD 420

Qy         387 RMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLS 446
               |||||||||||||||| || |||||||||||||||||||||||:||||||||||||||||
Db         421 RMKNSCIVCNMGHSNTETDVASLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLNLS 480

Qy         447 CSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTE 506
               ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db         481 CSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTE 540

Qy         507 LTDDQAKYLGLNKNGPFKPNYYRY 530
               |||:||||||||||||||||||||
Db         541 LTDEQAKYLGLNKNGPFKPNYYRY 564
```

RESULT 9
US-11-097-143-2781
; Sequence 2781, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2781
; LENGTH: 521
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-2781

Query Match 70.8%; Score 1970; DB 6; Length 521;
Best Local Similarity 77.6%; Pred. No. 1.5e-170;
Matches 370; Conservative 48; Mismatches 55; Indels 4; Gaps 3;

```
Qy          57 KTGRRSLSRSISQSSTDSYSSAASYTDSSD--DEVSPREKQQTNSKGSSNFCVKNI-KQA 113
               |   |  |||:| |||||:|| |||| ||:  |:| |||| | ||||||:|||:||  |
Db          46 KKTSRYRSRSLSASSTDSFSS-ASYTGSSEDGDDVPPREKVQKNSKGSSDFCVRNIGAQH 104

Qy         114 EFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCR 173
               |||||||||:|  :|:|:||| :||| |||||||| |||||||||  ||| |
Db         105 AFGRREIEIAEQEMPGIIALKKRAAEDKPLKDAKIVGCTHINAQTAVLIETLVELGASVR 164

Qy         174 WSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDPWWCIDRCVNMDGWQANMILDDGGDL 233
               |:|||||||||||||||:|: :|||:||:|:|||||||||| : || |||||||||
Db         165 WAACNIYSTQNEVAAALAESGIPIFAWRGETEEDFWWCIDRCVNAENWQPNMILDDGGDA 224

Qy         234 THWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCC 293
               || : |||| :|| ::|||||||||||||||||||||| |||||||||||| ||||||  |
Db         225 THLMLKKYPTMFKLVKGIVEESVTGVHRLYQLSKAGKLTVPAMNVNDSVTKTKFDNLYSC 284

Qy         294 RESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMD 353
               :||||| |||:|||||||||||||:||||| ||| | ||||||||||||| ||
Db         285 KESILDSLKRSTDVMFGGKQVVVCGYGDVGKGCAQALKGQGCIVYITEIDPICALQASMD 344

Qy         354 GFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPE 413
               |||||||||||| ||:|:| ||||||| |||:|:||: |||||||||||||||  ||||:
Db         345 GFRVVKLNEVIRNVDIVVTATGNKNVVVREHMDKMKSGCIVCNMGHSNTEIDVNGLRTPD 404

Qy         414 LTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAP 473
               ||||:|||||||:|||:|| ::|||||||:|||||::|:|: :|||: |||||||||:|||
Db         405 LTWEKVRSQVDHIIWPEGKYIILLAEGRLVNLSCSSIPSFAVSITSATQALALIELFNAP 464

Qy         474 EGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
               |||| ||||||||||||||||:||||||||:|:||||:|||| ||||||||||
Db         465 PGRYKSDVYLLPKKMDEYVASLHLPTFDAHLTELSDEQAKYMGLNKAGPFKPNYYRY 521
```

RESULT 10
US-11-097-143-41922
; Sequence 41922, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.

```
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41922
;   LENGTH: 504
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-11-097-143-41922

  Query Match             65.4%;  Score 1819.5;  DB 6;  Length 504;
  Best Local Similarity   71.5%;  Pred. No. 7.9e-157;
  Matches  334;  Conservative   64;  Mismatches   66;  Indels    3;  Gaps    3;

Qy         64 SRSISQSSTDSYSSAASYTDSSDDE-VSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEI 122
              :|  |  || |:||   :| |||:| |||::  | || | ::||||:| :: ||||||||
Db         27 ARRFSDVSTCSFSSTC-FTGSSDEEDVSPKDNHQRNSAGGTDFCVKSISKSAFGRREIEI 85

Qy        123 AEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYST 182
              || :|  :::|||||: |||| || ||||||: ||:|||||||  |||  ||:|||||||
Db         86 AESEMPGIMTLRKRAKDEKPLKGANIVGCTHVNAQSAVLIETLVQLGATVRWAACNIYST 145

Qy        183 QNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYP 242
              || |||||||||: :|||:||:|::||||:|| :  |||| |:||||||| || : ||||
Db        146 QNAVAAALAEAGIPIFAWRGETEEEFWWCLDRAIYSDGWQPNLILDDGGDATHLMLKKYP 205

Qy        243 NVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLK 302
              : || |||||||||||||||| ||| ||| |||:|||||||| |||  | ||:|||| ||
Db        206 DYFKAIRGIVEESVTGVHRLYMLSKGGKLTVPAINVNDSVTKNKFDTFYTCRDSILDSLK 265

Qy        303 RTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNE 362
              ||||:|||||||:||||:|||||  :||  | |||:||:|||||||| |||||||:|||
Db        266 RTTDIMFGGKQVVICGYGDVGKGCAQSLKGQGCIVYVTEVDPICALQAAMDGFRVVRLNE 325

Qy        363 VIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQ 422
              ||| ||||:| ||||||:||:|::|||| ||:|||||| :||||  | ||||||||||||
Db        326 VIRTVDVVVTATGNKNVITRDHMNRMKNGCILCNMGHSCSEIDVNGLHTPELTWERVRSQ 385

Qy        423 VDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVY 482
              |||: |||| ::|||||||:||||||: :||:|: ::|||||||||::|| |||| |||
Db        386 VDHIRWPDGRMIILLAEGRLVNLSCSTISSFVVSVASSTQALALIELFSAP-GRYKSDVY 444

Qy        483 LLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYR 529
              |||||||||||||| :||||||||||:|:|::|||| |||| ||||
Db        445 LLPKKMDEYVASLHLATFDAHLTELTDEQSKFMGLNKAGPFKANYYR 491


RESULT 11
US-10-450-763-42388
; Sequence 42388, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 42388
;   LENGTH: 455
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (356)..(388)
;   OTHER INFORMATION: S-adenosyl-L-homocysteine hydrolase proteins domain
;   OTHER INFORMATION: identified by eMATRIX, accession number BL00738F, p-value=7.261e
;   OTHER INFORMATION: 29, raw score of 12.23
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (186)..(447)
;   OTHER INFORMATION: S-adenosyl-L-homocysteine hydrolase domain identified by
;   OTHER INFORMATION: PFam, accession name AdoHcyase, E-value=6.2e-171, PFam score of
;   OTHER INFORMATION: 581.3
US-10-450-763-42388

  Query Match             57.6%;  Score 1601.5;  DB 5;  Length 455;
  Best Local Similarity   79.6%;  Pred. No. 5.5e-137;
  Matches  308;  Conservative   29;  Mismatches   23;  Indels   27;  Gaps    3;

Qy          8 PLPGVGE----------ELKQAKEIEDAEKY-------------SFMATVTKAPK--K 40
              |:|| |             :: ||  :: :: :             |   |||:||:  |
Db         61 PVPGPGSGPAAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 120

Qy         41 QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSK 100
              :|||||  ||| | ||| ||||||||||||||||||||||||||||||| |||:||| |||
Db        121 KIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSK 180

Qy        101 GSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAV 160
              |||:|||||||||||||||||:| ||::|||||||||||||||||||||||||||||||||
Db        181 GSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGAKIVGCTHITAQTAV 240

Qy        161 LIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDG 220
              |:||| |||||||:|||||||| |||||||||:|  |||||||||||||||||||::|
Db        241 LMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGESEDDFWWCIDRCVNVEG 300

Qy        221 WQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVND 280
              || |||||||||||:|||||:||||:|||||||||||||||||||||||||||||||||||
Db        301 WQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLSKAGKLCVPAMNVND 360

Qy        281 SVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYIT 340
              |||||||||||||||||||||:|||||||||||||||||||||||:|:|||:|
Db        361 SVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGCCAALKAMGSIVYVT 420

Qy        341 EIDPICALQACMDGFRVVKLNEVIRQV 367
              |||||||||||||:|||||||||
Db        421 EIDPICALQACMDGFRLVKLNEVIRQV 447


RESULT 12
US-10-369-493-10135
; Sequence 10135, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
```

```
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10135
;   LENGTH: 426
;   TYPE: PRT
;   ORGANISM: magnetite-containing magnetic coccus
US-10-369-493-10135

  Query Match             46.4%;  Score 1289;  DB 4;  Length 426;
  Best Local Similarity   56.4%;  Pred. No. 1.7e-108;
  Matches  242;  Conservative   81;  Mismatches  100;  Indels    6;  Gaps    5;

Qy        105 FCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIET 164
              | | :|  | :||:|||||| :|  |:|||     |:|| ||:: || |:| ||||||||
Db          1 FKVADIGLAAWGRKEIEIAETEMPGLVSLRSEYGAEQPLKGARVAGCLHMTIQTAVLIET 60

Qy        165 LCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCV-NMDGWQA 223
              | ||||: |||:|||:|||:  |||:| ||| |||||||:|::: ||::: :   |||
Db         61 LTALGAEVRWSSCNIFSTQDHAAAAIAAAGVPVFAWKGETEEEYVWCVEQTIKGPDGWVP 120

Qy        224 NMILDDGGDLTHWVYK-KYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSV 282
              |||||||||||  ::: :|  : | :||| ||: ||| |||:::: |||  || ||||||
Db        121 NMILDDGGDLTEILHRPEYVEIMKGVRGISEETTTGVLRLYEMAQKGKLECPAFNVNDSV 180

Qy        283 TKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEI 342
              || |||||| ||||::||:|| |||| || ||||||:|||||  | : |||:|::|||
Db        181 TKSKFDNLYGCRESLMDGIKRATDVMIAGKIGVVCGYGDVGKGCAQAFRGLGAVVWVTEI 240

Qy        343 DPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNT 402
              |||||||| |:|:||| : |   | |: :| |||:::|| |:::||:  ||||:|| ::
Db        241 DPICALQAAMEGYRVVTMEEAASQADIFVTATGNRDIITRAHMEQMKDQAIVCNIGHFDS 300

Qy        403 EIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATT 461
              |||: |:|  || |: :: |||||  | | |:::|||||:|| |:|  |:||:| : |
Db        301 EIDIASIR--ELKWDNIKPQVDHVTMPSGNRIIVLAEGRLVNLGCATGHPSFVMSNSFTN 358

Qy        462 QALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNG 521
              | :| |||:|  |  |:  ||:||| :|| || |||    | || ||  || |||:: :|
Db        359 QVMAQIELFNRGE-MYENRVYVLPKHLDEKVARLHLAKVGAKLTTLTQVQADYLGISVDG 417

Qy        522 PFKPNYYRY 530
              |:| ::|||
Db        418 PYKSDHYRY 426


RESULT 13
US-10-369-493-4783
; Sequence 4783, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4783
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Burkholderia fungorum
US-10-369-493-4783

  Query Match             42.7%;  Score 1187;  DB 4;  Length 461;
  Best Local Similarity   50.4%;  Pred. No. 3.9e-99;
  Matches  234;  Conservative   80;  Mismatches  110;  Indels   40;  Gaps    6;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163
              :| | ::  |::||:|: ||| :|  |:  |:  : ::|| ||:| |  |:| || ||||
Db          1 DFIVADMSLADWGRKELNIAETEMPGLVQTREEYKVQQPLKGARIAGSLHMTIQTGVLIE 60

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNM-DGWQ 222
              || ||||  ||::|||:|||:  |||:|:||  |||:|||| |::|    |     :|
Db         61 TLTALGADVRWASCNIFSTQDHAAAAIAKAGTPVFAFKGESLDEYWEFSHRIFEWPNGEF 120

Qy        223 ANMILDDGGDLTHW----------------------VYKK-----------YPNVFKK 247
              ||||||||||  |                      :||           |
Db        121 ANMILDDGGDATLLLILGSKAEKDRSVISKPTNEEEVALYKSIERHLDADPTWYSTRLAH 180

Qy        248 IRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDV 307
              |:|: ||: ||||||||: | |:|  ||:|||||||| |||||| ||||::||:|| |||
Db        181 IKGVTEETTTGVHRLYQMEKEGRLPFPAINVNDSVTKSKFDNLYGCRESLVDGIKRATDV 240

Qy        308 MFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQV 367
              |  ||  || |||:|||||  :|: ||| |::||||||||||| |:|:||| :     :
Db        241 MIAGKIAVVAGYGDVGKGCAQSLRGLGATVWVTEIDPICALQAAMEGYRVVTMEYAADKA 300

Qy        368 DVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVI 427
              |: :| ||| :|:  :|:  |::: ||||:|| ::|||| | |  :  |: :: ||||:|
Db        301 DIFVTATGNYHVIGHDHMKAMRHNAIVCNIGHFDSEIDVASTR--QYQWDNIKPQVDHII 358

Qy        428 WPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPK 486
              :||||||:|||||||:|| |:|  |:||:| : | | || |||:     :|:  ||:|||
Db        359 FPDGKRVILLAEGRLVNLGCATGHPSFVMSNSFTNQTLAQIELF-TQGNKYENKVYVLPK 417

Qy        487 KMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              :|| || |||    |:|| |:|:|| |:|::||||||:|||
Db        418 HLDEKVARLHLARIGANLTVLSDEQAGYIGVDKNGPFKPNHYRY 461


RESULT 14
US-10-369-493-7543
; Sequence 7543, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7543
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Burkholderia cepacia
US-10-369-493-7543

  Query Match             42.7%;  Score 1187;  DB 4;  Length 461;
  Best Local Similarity   50.4%;  Pred. No. 3.9e-99;
  Matches  234;  Conservative   80;  Mismatches  110;  Indels   40;  Gaps    6;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163
              :| | ::  |::||:|: ||| :|  |:  |:  : ::|| ||:| |  |:| || ||||
Db          1 DFIVADMSLADWGRKELNIAETEMPGLVQTREEYKVQQPLKGARIAGSLHMTIQTGVLIE 60

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNM-DGWQ 222
              || ||||  ||::|||:|||:  |||:|:||  |||:|||| |::|    |     :|
Db         61 TLTALGADVRWASCNIFSTQDHAAAAIAKAGTPVFAFKGESLDEYWEFSHRIFEWPNGEF 120

Qy        223 ANMILDDGGDLTHW----------------------VYKK-----------YPNVFKK 247
              ||||||||||  |                      :||           |
Db        121 ANMILDDGGDATLLLILGSKAEKDRSVISKPTNEEEVALYKSIERHLDADPTWYSTRLAH 180

Qy        248 IRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDV 307
              |:|: ||: ||||||||: | |:|  ||:|||||||| |||||| ||||::||:|| |||
Db        181 IKGVTEETTTGVHRLYQMEKEGRLPFPAINVNDSVTKSKFDNLYGCRESLVDGIKRATDV 240

Qy        308 MFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQV 367
              |  ||  || |||:|||||  :|: ||| |::||||||||||| |:|:||| :     :
Db        241 MIAGKIAVVAGYGDVGKGCAQSLRGLGATVWVTEIDPICALQAAMEGYRVVTMEYAADKA 300

Qy        368 DVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVI 427
              |: :| ||| :|:  :|:  |::: ||||:|| ::|||| | |  :  |: :: ||||:|
Db        301 DIFVTATGNYHVIGHDHMKAMRHNAIVCNIGHFDSEIDVASTR--QYQWDNIKPQVDHII 358

Qy        428 WPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPK 486
              :||||||:|||||||:|| |:|  |:||:| : | | || |||:     :|:  ||:|||
Db        359 FPDGKRVILLAEGRLVNLGCATGHPSFVMSNSFTNQTLAQIELF-TQGNKYENKVYVLPK 417

Qy        487 KMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              :|| || |||    |:|| |:|:|| |:|::||||||:|||
Db        418 HLDEKVARLHLARIGANLTVLSDEQAGYIGVDKNGPFKPNHYRY 461


Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163
              :| | ::  |::||:|: ||| :|  |:  |:  : ::|| ||:| |  |:| || ||||
Db          1 DFIVADMSLADWGRKELNIAETEMPGLVQTREEYKVQQPLKGARIAGSLHMTIQTGVLIE 60

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNM-DGWQ 222
              || ||||  ||::|||:|||:  |||:|:||  |||:|||| |::|    |     :|
```

```
Db         61 TLTALGADVRWASCNIFSTQDHAAAAIAKAGTPVFAFKGESLDEYWEFSHRIFEWPNGEF 120

Qy        223 ANMILDDGGDLTHW------------------------VYKK-----------YPNVFKK 247
              |||||||||| |                          :||           |
Db        121 ANMILDDGGDATLLLILGSKAEKDRSVISKPTNEEEVALYKSIERHLDADPTWYSTRLAH 180

Qy        248 IRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDV 307
              |:|: ||: |||||||| | |:|  ||:|||||||| |||||| ||||::||:|| |||
Db        181 IKGVTEETTTGVHRLYQMEKEGRLPFPAINVNDSVTKSKFDNLYGCRESLVDGIKRATDV 240

Qy        308 MFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQV 367
              |  ||  || |||:|||||  :|: ||| |::|||||||||| |:|:||| :     :
Db        241 MIAGKIAVVAGYGDVGKGCAQSLRGLGATVWVTEIDPICALQAAMEGYRVVTMEYAADKA 300

Qy        368 DVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVI 427
              |: :| ||| :|:  :|:  |::: ||||:|| ::|||| | |  :  |: :: ||||:|
Db        301 DIFVTATGNYHVIGHDHMKAMRHNAIVCNIGHFDSEIDVASTR--QYQWDNIKPQVDHII 358

Qy        428 WPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPK 486
              :||||||:|||||||:|| |:|  |:||:| : | | || |||:     :|:  ||:|||
Db        359 FPDGKRVILLAEGRLVNLGCATGHPSFVMSNSFTNQTLAQIELF-TQGNKYENKVYVLPK 417

Qy        487 KMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
               :|| || |||    |:|| |:|:|| |:|::||||||||:|||
Db        418 HLDEKVARLHLARIGANLTVLSDEQAGYIGVDKNGPFKPNHYRY 461


RESULT 15
US-10-369-493-19704
; Sequence 19704, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19704
;   LENGTH: 463
;   TYPE: PRT
;   ORGANISM: Nitrosomonas europaea
US-10-369-493-19704

  Query Match             42.6%;  Score 1185;  DB 4;  Length 463;
  Best Local Similarity   49.3%;  Pred. No. 6e-99;
  Matches  233;  Conservative   81;  Mismatches   97;  Indels   62;  Gaps    8;

Qy        107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
              | ::  |::||:|| ||| :|  |::||::   :||||||:| |  |:| |||||||||
Db          4 VADLSLADWGRKEIAIAETEMPGLMALREQYADKKPLAGARIAGSLHMTIQTAVLIETLV 63

Qy        167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFW--------WCIDRCVNM 218
              ||||: ||::|||:|||:  |||:|   : |||:|||| :::|        |  |
Db         64 ALGAEVRWASCNIFSTQDHAAAAIAARNIPVFAYKGESLEEYWDYAHQIFEWASD----- 118

Qy        219 DGWQANMILDDGGDLT-------------------------------------HWVYK 239
                  |||||||||| |                                     |
Db        119 GSHTANMILDDGGDATLLLILGSKAERDPSVIANPTNEEEQVLFASIRSRLASHPGW--- 175

Qy        240 KYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILD 299
               |     |||: ||: |||||||:: | |:|  ||:|||||||| |||||| ||||::|
Db        176 -YSRNLAAIRGVTEETTTGVHRLYEMEKKGELPFPAINVNDSVTKSKFDNLYGCRESLVD 234

Qy        300 GLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVK 359
              |:|| ||||  ||  ||||||:|||||  :|: ||| |:|||||||||||| |:|:|||

Db        235 GIKRATDVMIAGKIAVVCGYGDVGKGCAQSLRGLGATVWITEIDPICALQAAMEGYRVVT 294

Qy        360 LNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERV 419
              :::   : |: :| |||  |:| :|: :||:  |:||:|| ::|||: |::  :  || :
Db        295 MDDACDKADIFVTATGNLRVITHDHMLKMKDQSIICNIGHFDSEIDIASVQ--KYQWENI 352

Qy        420 RSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGR-Y 477
              : ||||||:| |:|:::||:|||:|| |:|  |:||:| : | || |||:     |: |
Db        353 KPQVDHVIFPTGRRIIVLAQGRLVNLGCATGHPSFVMSSSFTNQVLAQIELWQ--NGKDY 410

Qy        478 KQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              :: ||:|||::|| || |||      ||||||:|| || |:||||:||  |||
Db        411 QKKVYVLPKRLDEMVARLHLGKLGVKLTELTDEQAHYLNLDKNGPYKPEMYRY 463


Search completed: December 14, 2005, 08:31:26
Job time : 123.215 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:13:48 ; Search time 37.6183 Seconds
(without alignments)
1164.808 Million cell updates/sec

Title: US-10-683-610-1
Perfect score: 2781
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....QAKYLGLNKNGPFKPNYYRY 530

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2618	94.1	500	1	US-08-896-005-1	Sequence 1, Appli
2	1819.5	65.4	504	1	US-08-896-005-3	Sequence 3, Appli
3	1160	41.7	432	1	US-08-896-005-5	Sequence 5, Appli
4	1151	41.4	432	1	US-08-896-005-4	Sequence 4, Appli
5	1151	41.4	432	2	US-09-347-878-1	Sequence 1, Appli
6	1151	41.4	432	2	US-09-546-013-3	Sequence 3, Appli
7	1150	41.4	447	2	US-09-949-016-10359	Sequence 10359, A
8	1108	39.8	515	2	US-09-248-796A-18264	Sequence 18264, A
9	1099.5	39.5	476	2	US-09-902-540-10728	Sequence 10728, A
10	1086	39.1	485	2	US-08-930-894-2	Sequence 2, Appli
11	1079.5	38.8	502	2	US-09-252-991A-23501	Sequence 23501, A
12	1074	38.6	498	2	US-09-919-854-2	Sequence 2, Appli
13	1062	38.2	463	2	US-09-328-352-4600	Sequence 4600, Ap
14	1032	37.1	486	2	US-09-759-990A-7	Sequence 7, Appli
15	409.5	14.7	387	2	US-09-489-039A-8393	Sequence 8393, Ap
16	322	11.6	68	2	US-09-270-767-33605	Sequence 33605, A
17	322	11.6	68	2	US-09-270-767-48822	Sequence 48822, A
18	313	11.3	99	2	US-08-930-894-8	Sequence 8, Appli
19	216	7.8	139	2	US-08-930-894-6	Sequence 6, Appli
20	211	7.6	97	1	US-08-204-740-7	Sequence 7, Appli
21	211	7.6	97	2	US-09-081-167A-7	Sequence 7, Appli
22	211	7.6	97	2	US-09-081-395-7	Sequence 7, Appli
23	211	7.6	97	2	US-09-416-833-7	Sequence 7, Appli
24	211	7.6	97	4	PCT-US95-02521-7	Sequence 7, Appli
25	209.5	7.5	138	2	US-08-930-894-5	Sequence 5, Appli
26	207.5	7.5	140	2	US-08-930-894-4	Sequence 4, Appli
27	204	7.3	81	2	US-09-513-999C-5999	Sequence 5999, Ap
28	196.5	7.1	95	1	US-08-204-740-9	Sequence 9, Appli
29	196.5	7.1	95	2	US-09-081-167A-9	Sequence 9, Appli
30	196.5	7.1	95	2	US-09-081-395-9	Sequence 9, Appli
31	196.5	7.1	95	2	US-09-416-833-9	Sequence 9, Appli
32	196.5	7.1	95	4	PCT-US95-02521-9	Sequence 9, Appli
33	186	6.7	139	2	US-08-930-894-7	Sequence 7, Appli
34	112.5	4.0	416	2	US-09-543-681A-5020	Sequence 5020, Ap
35	106	3.8	1481	2	US-09-231-899-70	Sequence 70, Appl
36	105.5	3.8	539	2	US-09-489-039A-13875	Sequence 13875, A
37	105	3.8	760	2	US-09-489-039A-9113	Sequence 9113, Ap
38	101	3.6	783	2	US-09-540-236-2284	Sequence 2284, Ap
39	97	3.5	542	2	US-09-538-092-289	Sequence 289, App
40	97	3.5	1078	2	US-09-248-796A-20284	Sequence 20284, A
41	97	3.5	1409	2	US-09-949-001-22	Sequence 22, Appl
42	96.5	3.5	1432	2	US-08-781-891-71	Sequence 71, Appl
43	96.5	3.5	1432	2	US-09-618-166-71	Sequence 71, Appl
44	96.5	3.5	1432	2	US-09-949-001-16	Sequence 16, Appl
45	95.5	3.4	374	2	US-09-647-224A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-896-005-1
; Sequence 1, Application US/08896005
; Patent No. 5854023
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,005
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0337 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADTUT04
; CLONE: 1519044
US-08-896-005-1

Query Match 94.1%; Score 2618; DB 1; Length 500;
Best Local Similarity 99.8%; Pred. No. 1.2e-287;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         31 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 60

Qy         91 PREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVG 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVG 120

Qy        151 CTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWW 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWW 180

Qy        211 CIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGK 270
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        181 CIDRCVNMDGWQANMILDDGGDLTHWVCKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGK 240

Qy        271 LCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAAL 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAAL 300

Qy        331 KALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKN 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKN 360

Qy        391 SCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTV 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTV 420

Qy        451 PTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDD 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 PTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDD 480

Qy        511 QAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||||||
Db        481 QAKYLGLNKNGPFKPNYYRY 500


RESULT 2
US-08-896-005-3
; Sequence 3, Application US/08896005
; Patent No. 5854023
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/896,005
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0337 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  504 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  969078
US-08-896-005-3

  Query Match             65.4%;  Score 1819.5;  DB 1;  Length 504;
  Best Local Similarity   71.5%;  Pred. No. 5.4e-197;
  Matches  334;  Conservative   64;  Mismatches   66;  Indels    3;  Gaps    3;

Qy         64 SRSISQSSTDSYSSAASYTDSSDDE-VSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEI 122
              :|  |  || |:||   :| |||:| |||::  | || | ::||||:| :: |||||||| 
Db         27 ARRFSDVSTCSFSSTC-FTGSSDEEDVSPKDNHQRNSAGGTDFCVKSISKSAFGRREIEI 85

Qy        123 AEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYST 182
              || :|  :::|||||: |||| || ||||||: ||:|||||||  |||  ||:|||||||
Db         86 AESEMPGIMTLRKRAKDEKPLKGANIVGCTHVNAQSAVLIETLVQLGATVRWAACNIYST 145

Qy        183 QNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYP 242
              || |||||||||: :|||:||:|::||||:|| :  |||| |:||||||| || : ||||
Db        146 QNAVAAALAEAGIPIFAWRGETEEEFWWCLDRAIYSDGWQPNLILDDGGDATHLMLKKYP 205

Qy        243 NVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLK 302
              : || |||||||||||||||| ||| ||| |||:|||||||| |||  | ||:|||| ||
Db        206 DYFKAIRGIVEESVTGVHRLYMLSKGGKLTVPAINVNDSVTKNKFDTFYTCRDSILDSLK 265

Qy        303 RTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNE 362
              ||||:|||||||:||||:|||||  :||  | |||:||:|||||||| |||||||:|||
Db        266 RTTDIMFGGKQVVICGYGDVGKGCAQSLKGQGCIVYVTEVDPICALQAAMDGFRVVRLNE 325

Qy        363 VIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQ 422
              ||| ||||:| |||||:||:|::|||| ||:|||||| :||||  | ||||||||||||
Db        326 VIRTVDVVVTATGNKNVITRDHMNRMKNGCILCNMGHSCSEIDVNGLHTPELTWERVRSQ 385

Qy        423 VDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVY 482
              |||: |||: ::|||||||:|||||: :||:|: ::|||||||||::|| |||| |||
Db        386 VDHIRWPDGRMIILLAEGRLVNLSCSTISSFVVSVASSTQALALIELFSAP-GRYKSDVY 444

Qy        483 LLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYR 529
              |||||||||||||| :|||||||||:|:|::|||| |||| ||||
Db        445 LLPKKMDEYVASLHLATFDAHLTELTDEQSKFMGLNKAGPFKANYYR 491


RESULT 3
US-08-896-005-5
; Sequence 5, Application US/08896005
; Patent No. 5854023
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/896,005
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0337 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  432 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  904132
US-08-896-005-5

  Query Match             41.7%;  Score 1160;  DB 1;  Length 432;
  Best Local Similarity   51.8%;  Pred. No. 3.1e-122;
  Matches  220;  Conservative   83;  Mismatches  120;  Indels    2;  Gaps    2;

Qy        107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
              | :|  | :||: ::||| :|  |: :|:    ||| ||:| || |:| :||||||||
Db          9 VADIGLAAWGRKALDIAENEMPGLMRMREMYSASKPLKGARIAGCLHMTVETAVLIETLV 68

Qy        167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMI 226
              ||||: |||:|||:|||:  |||:|:||: |||||||:::::: |||:: ::      |||
Db         69 ALGAEVRWSSCNIFSTQDHAAAAIAKAGIPVFAWKGETDEEYLWCIEQTLHFKDGPLNMI 128

Qy        227 LDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQK 286
              ||||||||: :: ||| :   |||| ||: |||| ||::   | | |||:||||||| |
Db        129 LDDGGDLTNLIHTKYPQLLSGIRGISEETTTGVHNLYKMMSNGILNVPAINVNDSVTKSK 188

Qy        287 FDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPIC 346
              ||||| |||||::||:|| ||||  ||  || |||:|||||  ||:  || | |||||||
Db        189 FDNLYGCRESLIDGIKRATDVMIAGKVAVVAGYGDVGKGCAQALRGFGARVIITEIDPIN 248

Qy        347 ALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDV 406
              |||| |:|: |  ::|  :: :: :| ||  :::   | ::||:  ||||:|| : ||||
Db        249 ALQAAMEGYEVTTMDEACKEGNIFVTTTGCVDIILGRHFEQMKDDAIVCNIGHFDVEIDV 308

Qy        407 TSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALA 465
              |    :    :: |||     :|:|::|||||||:|| |:   |:||:| : | | :|
Db        309 KWLNENAVEKVNIKPQVDRYWLKNGRRIILLAEGRLVNLGCAMGHPSFVMSNSFTNQVMA 368

Qy        466 LIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKP 525
              |||:  |: :|   |: ||||:|| ||  ||   :  ||:||: ||:|||:  ||||||
Db        369 QIELWTHPD-KYPVGVHFLPKKLDEAVAEAHLGKLNVKLTKLTEKQAQYLGMPINGPFKP 427

Qy        526 NYYRY 530
              ::|||
Db        428 DHYRY 432


RESULT 4
US-08-896-005-4
; Sequence 4, Application US/08896005
; Patent No. 5854023
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/896,005
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0337 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  432 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  178277
US-08-896-005-4

  Query Match             41.4%;  Score 1151;  DB 1;  Length 432;
  Best Local Similarity   51.3%;  Pred. No. 3.3e-121;
  Matches  218;  Conservative   85;  Mismatches  120;  Indels    2;  Gaps    2;

Qy        107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
              | :|  | :||: ::||| :|  |: :|:|    ||| ||:| || |:| :||||||||
Db          9 VADIGLAAWGRKALDIAENEMPGLMRMRERYSASKPLKGARIAGCLHMTVETAVLIETLV 68

Qy        167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMI 226
              |||: :||:|||:|||| |||:|:||: |:|||||:::: |||:: :      |||
Db         69 TLGAEVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLWCIEQTLYFKDGPLNMI 128

Qy        227 LDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQK 286
              |||||||: :: ||| :   |||| ||: |||| ||::   | | |||:||||||| |
Db        129 LDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVHNLYKMMANGILKVPAINVNDSVTKSK 188

Qy        287 FDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPIC 346
              ||||| |||||::||:|| ||||  ||  || |||:|||||  ||:  || | |||||||
Db        189 FDNLYGCRESLIDGIKRATDVMIAGKVAVVAGYGDVGKGCAQALRGFGARVIITEIDPIN 248

Qy        347 ALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDV 406
              |||| |:|: |  ::|  :: :: :| ||  :::   | ::||:  ||||:|| : ||||
Db        249 ALQAAMEGYEVTTMDEACQEGNIFVTTTGCIDIILGRHFEQMKDDAIVCNIGHFDVEIDV 308

Qy        407 TSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALA 465
              |    :    :: |||     :|:|::|||||||:|| |:   |:||:| : | | :|
Db        309 KWLNENAVEKVNIKPQVDRYRLKNGRRIILLAEGRLVNLGCAMGHPSFVMSNSFTNQVMA 368

Qy        466 LIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKP 525
              |||:  |: :|   |: ||||:|| ||  ||   :  ||:||: ||:|||:: :|||||
Db        369 QIELWTHPD-KYPVGVHFLPKKLDEAVAEAHLGKLNVKLTKLTEKQAQYLGMSCDGPFKP 427

Qy        526 NYYRY 530
              ::|||
Db        428 DHYRY 432


RESULT 5
US-09-347-878-1
```

; Sequence 1, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase protein
; OTHER INFORMATION: sequence
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M61831/GenBank
US-09-347-878-1

Query Match 41.4%; Score 1151; DB 2; Length 432;
Best Local Similarity 51.3%; Pred. No. 3.3e-121;
Matches 218; Conservative 85; Mismatches 120; Indels 2; Gaps 2;

Qy 107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
Db 9 VADIGLAAWGRKALDIAENEMPGLMRMRERYSASKPLKGARIAGCLHMTVETAVLIETLV 68

Qy 167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMI 226
Db 69 TLGAEVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLWCIEQTLYFKDGPLNMI 128

Qy 227 LDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQK 286
Db 129 LDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVHNLYKMMANGILKVPAINVNDSVTKSK 188

Qy 287 FDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPIC 346
Db 189 FDNLYGCRESLIDGIKRATDVMIAGKVAVVAGYGDVGKGCAQALRGFGARVIITEIDPIN 248

Qy 347 ALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDV 406
Db 249 ALQAAMEGYEVTTMDEACQEGNIFVTTTGCIDIILGRHFEQMKDDAIVCNIGHFDVEIDV 308

Qy 407 TSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALA 465
Db 309 KWLNENAVEKVNIKPQVDRYRLKNGRRIILLAEGRLVNLGCAMGHPSFVMSNSFTNQVMA 368

Qy 466 LIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKP 525
Db 369 QIELWTHPD-KYPVGVHFLPKKLDEAVAEAHLGKLNVKLTKLTEKQAQYLGMSCDGPFKP 427

Qy 526 NYYRY 530
Db 428 DHYRY 432

RESULT 6
US-09-546-013-3
; Sequence 3, Application US/09546013
; Patent No. 6610504
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Shen
; TITLE OF INVENTION: METHODS FOR ASSAYING S-ADENOSYLMETHIONINE-DEPENDENT Methyltransfe
; FILE REFERENCE: 10937-1652
; CURRENT APPLICATION NUMBER: US/09/546,013
; CURRENT FILING DATE: 2000-04-10
; EARLIER APPLICATION NUMBER: 09/347,878
; EARLIER FILING DATE: 1999-07-06
; EARLIER APPLICATION NUMBER: 09/457,205
; EARLIER FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase protein
; OTHER INFORMATION: sequence
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M61831/GenBank
US-09-546-013-3

Query Match 41.4%; Score 1151; DB 2; Length 432;
Best Local Similarity 51.3%; Pred. No. 3.3e-121;
Matches 218; Conservative 85; Mismatches 120; Indels 2; Gaps 2;

Qy 107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
Db 9 VADIGLAAWGRKALDIAENEMPGLMRMRERYSASKPLKGARIAGCLHMTVETAVLIETLV 68

Qy 167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMI 226
Db 69 TLGAEVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLWCIEQTLYFKDGPLNMI 128

Qy 227 LDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQK 286
Db 129 LDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVHNLYKMMANGILKVPAINVNDSVTKSK 188

Qy 287 FDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPIC 346
Db 189 FDNLYGCRESLIDGIKRATDVMIAGKVAVVAGYGDVGKGCAQALRGFGARVIITEIDPIN 248

Qy 347 ALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDV 406
Db 249 ALQAAMEGYEVTTMDEACQEGNIFVTTTGCIDIILGRHFEQMKDDAIVCNIGHFDVEIDV 308

Qy 407 TSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALA 465
Db 309 KWLNENAVEKVNIKPQVDRYRLKNGRRIILLAEGRLVNLGCAMGHPSFVMSNSFTNQVMA 368

Qy 466 LIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKP 525
Db 369 QIELWTHPD-KYPVGVHFLPKKLDEAVAEAHLGKLNVKLTKLTEKQAQYLGMSCDGPFKP 427

Qy 526 NYYRY 530
Db 428 DHYRY 432

RESULT 7
US-09-949-016-10359
; Sequence 10359, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10359
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10359

```
  Query Match             41.4%;  Score 1150;  DB 2;  Length 447;
  Best Local Similarity   48.7%;  Pred. No. 4.6e-121;
  Matches  223;  Conservative   89;  Mismatches  126;  Indels   20;  Gaps    3;

Qy         74 SYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEPGRREIEIAEQDMSALISL 133
              | :|||| :|                      : | :|  | :||: ::||| :|  |: :
Db          9 SITSAASMSDKLP-----------------YKVADIGLAAWGRKALDIAENEMPGLMRM 50

Qy        134 RKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEA 193
              |:|    ||| ||:| || |:| :|||||||  |||: :||:|||:|||:  |||:|:|
Db         51 RERYSASKPLKGARIAGCLHMTVETAVLIETLVTLGAEVQWSSCNIFSTQDHAAAAIAKA 110

Qy        194 GVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVE 253
              |: |:|||||:::: |||:: :       ||||||||||: :: ||| :   |||| |
Db        111 GIPVYAWKGETDEEYLWCIEQTLYFKDGPLNMILDDGGDLTNLIHTKYPQLLPGIRGISE 170

Qy        254 ESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQ 313
              |: |||| ||::   | | |||:|||||||| |||||| ||||::||:|| ||||  ||
Db        171 ETTTGVHNLYKMMANGILKVPAINVNDSVTKSKFDNLYGCRESLIDGIKRATDVMIAGKV 230

Qy        314 VVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITC 373
               || |||:|||||  ||:  || | ||||||| |||| |:|: |  ::|  :: :: :|
Db        231 AVVAGYGDVGKGCAQALRGFGARVIITEIDPINALQAAMEGYEVTTMDEACQEGNIFVTT 290

Qy        374 TGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKR 433
              ||  :::   | ::||:  ||||:|| : ||||  |    :   :: |||     :|:|
Db        291 TGCIDIILGRHFEQMKDDAIVCNIGHFDVEIDVKWLNENAVEKVNIKPQVDRYRLKNGRR 350

Qy        434 VVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYV 492
              ::|||||||:|| |:   |:||:| : | | :| |||:  |: :|   |: ||||:|| |
Db        351 IILLAEGRLVNLGCAMGHPSFVMSNSFTNQVMAQIELWTHPD-KYPVGVHFLPKKLDEAV 409

Qy        493 ASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              |  ||   :  ||:||: ||:|||:: :|||||::|||
Db        410 AEAHLGKLNVKLTKLTEKQAQYLGMSCDGPFKPDHYRY 447


RESULT 8
US-09-248-796A-18264
; Sequence 18264, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18264
;   LENGTH: 515
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-18264

  Query Match             39.8%;  Score 1108;  DB 2;  Length 515;
  Best Local Similarity   49.0%;  Pred. No. 3.4e-116;
  Matches  220;  Conservative   86;  Mismatches  127;  Indels   16;  Gaps    4;

Qy         98 NSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQ 157
              :|  ::|: | :|  | |||::||::| :|  |: :||:    :|| ||:| || |:| |
Db         67 SSAPATNYKVADISLAAFGRKDIELSENEMPGLMYIRKKYGPSQPLKGARIAGCLHMTIQ 126

Qy        158 TAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDR--C 215
              |||||||| ||||:  ||:|||:|||:  |||:| ||| |||||||:|::: |||::
Db        127 TAVLIETLVALGAEVTWSSCNIFSTQDHAAAAIAAAGVPVFAWKGETEEEYQWCIEQQLF 186

Qy        216 VNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPA 275
               || : |:|||||||||  |::||| : :   |: ||: |||| ||:  : |||  ||  :
Db        187 AFKDGKKLNLILDDGGDLTSLVHEKYPEMLEDCYGLSEETTTGVHHLYKSLRDGKLKVPA 246

Qy        276 MNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGA 335
              :|||||||| |||||| ||||::||:|| ||||  ||  :|  |:|:|||||  ||  :||
Db        247 INVNDSVTKSKFDNLYGCRESLVDGIKRATDVMIAGKVAIVAGFGDVGKGCAMALHGMGA 306

Qy        336 IVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVC 395
              | :|||||| |||| : |::|  ::||     : :| || :::::| :| ::|    |||
Db        307 RVIVTEIDPINALQAAVSGYQVAPMDEVASIGQIFVTTTGCRDIITGKHFEQMPEHAIVC 366

Qy        396 NMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFV 454
              |:|| : |||| |      :  :: |||  :  :|: |:|||:|||:|| |:|   :||
Db        367 NIGHFDIEIDVAWLIANAESVVNIKPQVDRYLMKNGRHVILLADGRLVNLGCATGHSSFV 426

Qy        455 LSITATTQALALIELYNA---------PE----GRYKQDVYLLPKKMDEYVASLHLPSFD 501
              :| : : | || | |:||          ||     | :    |:|||| :|| ||  ||
Db        427 MSCSFSNQVLAQIALFNADNKEFREKFPEFAKTGPFDVGVHLLPKVLDETVARCHLDHLG 486

Qy        502 AHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              | || ||: ||:|||: :  |:| : |||
Db        487 AKLTTLTETQAEYLGIPEESPYKADIYRY 515


RESULT 9
US-09-902-540-10728
; Sequence 10728, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10728
;   LENGTH: 476
;   TYPE: PRT
;   ORGANISM: Myxococcus xanthus
US-09-902-540-10728

  Query Match             39.5%;  Score 1099.5;  DB 2;  Length 476;
  Best Local Similarity   47.0%;  Pred. No. 2.8e-115;
  Matches  220;  Conservative   82;  Mismatches  123;  Indels   43;  Gaps    5;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163
              :: : ::  |::||:|| ||| :| ||:::|:     :|| ||:: |  |:| |||||:|
Db         11 DYAIADLSLADWGRKEIRIAESEMPALMAIREEYAKTQPLKGARVTGSLHMTIQTAVLVE 70

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNM----- 218
              || ||||: ||::|||:|||:  |||| :||  ||| ||||  ::|    | :
Db         71 TLQALGAEVRWASCNIFSTQDHAAAALVQAGTPVFAHKGESLKEYWDFTHRIFDFGPAGS 130

Qy        219 DGWQANMILDDGGDLTHWVY--KK-------------------------------YPN 243
              |    |||||||||  |  ::  |:                               |
Db        131 DHEGPNMILDDGGDATLLMHLGKRAEKDASVLANPQSEEERELYAAIKAKLAEDATWYTR 190

Qy        244 VFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKR 303
              || |: ||: |||||| ::|  | |   |:||||||| |||||| ||||::||:|)
Db        191 KSAKILGVTEETTTGVHRLQEMSAKGTLLFRAINVNDSVTKSKFDNLYGCRESLVDGIKR 250

Qy        304 TTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEV 363
               ||||  ||  || |||:||||   ||:|| | |::||||||||||| |:|:||| ::
Db        251 ATDVMVAGKIAVVAGYGDVGKGSAQALRALSAQVWVTEIDPICALQAAMEGYRVVTMDYA 310

Qy        364 IRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQV 423
              : |: :| |||| |:| :|: :||:  ||||:|| : ||:| ||   :  || :: ||
Db        311 ADKADIFVTATGNKGVITHDHMAKMKDQAIVCNIGHFDNEIEVASLE--KYKWEEIKPQV 368
```

```
Qy        424 DHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGRYKQDVY 482
Db        369 DHIIFPDNKRIILLAKGRLVNLGCGTGHPSYVMSSSFANQTIAQIELYSHSANYEVGKVY 428

Qy        483 LLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
Db        429 VLPKHLDEKVARLQLKKLNAQLTELSQEQADYIGVKTSGPYKPDTYRY 476


RESULT 10
US-08-930-894-2
; Sequence 2, Application US/08930894
; Patent No. 6037524
;  GENERAL INFORMATION:
;    APPLICANT:  GREENLAND, Andrew James
;    APPLICANT:  DRAPER, John
;    APPLICANT:  SKIPSEY, Marc
;    APPLICANT:  WARNER, Simon
;    TITLE OF INVENTION:  S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pillsbury Madison & Sutro
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3918
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/930,894
;      FILING DATE:  09-OCT-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB96/00882
;      FILING DATE:  10-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9507381.3
;      FILING DATE:  10-APR-1995
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 485 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-930-894-2

  Query Match             39.1%;  Score 1086;  DB 2;  Length 485;
  Best Local Similarity   45.7%;  Pred. No. 9.7e-114;
  Matches  215;  Conservative   84;  Mismatches  125;  Indels   46;  Gaps    6;

Qy        107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
Db         16 VKDMSQADFGRLEIELAEVEMPGLMACRAEFGPAQPFKGAKITGSLHMTIQTAVLIETLT 75

Qy        167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNM-DGWQANM 225
Db         76 ALGPEVRWCSCNIFSTQDHAAAAIARDSASVFAWKGETLQEYWWCTERALDWGPGGGPDL 135

Qy        226 ILDDGGDLTHWVYK----------------------------------------KYPNV 244
Db        136 IVDDGGDTTLLIHEGVKAEEEYEKTGKMPDPASTDNAEFQIVLTIIRDGLKVDPTKYRKM 195

Qy        245 FKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRT 304
Db        196 KDRIVGVSEETTTGVKRLYQMQANNSLLFPAINVNDSVTKSKFDNLYGCRHSLPDGLMRA 255

Qy        305 TDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVI 364
Db        256 TDVMIAGKVAVVCGYGDVGEGCAAALKQAGARVIVTEIDPICALQALMEGLQVLTLEDVV 315

Qy        365 RQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRT-PELTWERVRSQV 423
Db        316 SEADIFVTTTGNKDIIMLDHMRKMKNNAIVCNIGHFDNEIDMLGLETYPGIKRITIKPQT 375

Qy        424 DHVIWPD-GKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYN-APEGRYKQD 480
Db        376 DRWVFPETNTGIIVLAEGRLMNLGCATGHPSFVMSCSFTNQVIAQLELWNEKASGKYEKK 435

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
Db        436 VYVLPKHLDEKVAALHLGKLGAKLTKLSPSQADYISVPIEGPYKPPHYRY 485


RESULT 11
US-09-252-991A-23501
; Sequence 23501, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 23501
;   LENGTH: 502
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23501

  Query Match             38.8%;  Score 1079.5;  DB 2;  Length 502;
  Best Local Similarity   45.1%;  Pred. No. 5.7e-113;
  Matches  225;  Conservative   93;  Mismatches  144;  Indels   37;  Gaps    9;

Qy         65 RSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAE 124
Db          8 RALRPASPVCQRTNGAHSQTTNGEL--RMSAVMTPAGFTDYKVADITLAAWGRRELIIAE 65

Qy        125 QDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQN 184
Db         66 SEMPALMGLRRKYAGQQPLKGAKILGCIHMTIQTGVLIETLVALGAEVRWSSCNIFSTQD 125

Qy        185 EVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDG--WQANMILDDGGDLTHWVYKKYP 242
Db        126 QAAAAIAAAGIPVFAWKGETEEEYEWCIEQTILKDGQPWDANMVLDDGGDLTEILHKKYP 185

Qy        243 NVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLK 302
Db        186 QMLERIHGITEETTTGVHRLLDMLKNGALKVPAINVNDSVTKSKNDNKYGCRHSLNDAIK 245

Qy        303 RTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVK--- 359
Db        246 RGTDHLLSGKQALVIGYGDVGKGSSQSLRQEGMIVKVAEVDPICAMQACMDGFEVVSPYK 305

Qy        360 --LNE---------VIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTS 408
Db        306 NGINDGTEASIDAALLGKIDLIVTTTGNVNVCDANMLKALKKRAVVCNIGHFDNEIDTAF 365

Qy        409 LRTPELTWERVRSQVDHVIWPDGK---------RVVLLAEGRLLNLSCST-VPTFVLSIT 458
Db        366 MR-KNWAWEEVKPQV-HKIHRTGKDGFDAHNDDYLILLAEGRLVNLGNATGHPSRIMDGS 423

Qy        459 ATTQALALIELYN-------APEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQ 511
Db        424 FANQVLAQIHLFEQKYADLPAAEKAKRLSVEVLPKKLDEEVALEMVKGFGGVVTQLTPKQ 483
```

```
Qy        512 AKYLGLNKNGPFKPNYYRY 530
              |:|:|::  |||||: |||
Db        484 AEYIGVSVEGPFKPDTYRY 502


RESULT 12
US-09-919-854-2
; Sequence 2, Application US/09919854
; Patent No. 6759224
; GENERAL INFORMATION:
;  APPLICANT: FARWICK, MIKE
;  APPLICANT:  HUTHMACHER, KLAUS
;  APPLICANT:  BREHME, JENNIFER
;  APPLICANT:  PFEFFERLE, WALTER
;  APPLICANT:  BINDER, MICHAEL
;  APPLICANT:  GREISSINGER, DIETER
;  APPLICANT:  THIERBACH, GEORG
;  TITLE OF INVENTION: Nucleotide sequences which code for the sahH gene
;  FILE REFERENCE: 211739US0X
;  CURRENT APPLICATION NUMBER: US/09/919,854
;  CURRENT FILING DATE:  2002-12-30
;  PRIOR APPLICATION NUMBER: DE 100 44 706.6
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: DE 101 09 685.2
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: US 60/294,277
;  PRIOR FILING DATE: 2001-05-31
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 498
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-919-854-2

  Query Match             38.6%;  Score 1074;  DB 2;  Length 498;
  Best Local Similarity   45.4%;  Pred. No. 2.4e-112;
  Matches  215;  Conservative   86;  Mismatches  125;  Indels   48;  Gaps    5;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163
              :| | ::  || || :| :|| :|  |: |||    |:|| ||:| |  |:| |||||||
Db         26 DFKVADLSLAEAGRHQIRLAEYEMPGLMQLRKEFADEQPLKGARIAGSIHMTVQTAVLIE 85

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALA--------EAGVAVFAWKGESEDDFWWCIDRC 215
              || ||||: ||::|||:|||:| |||:           ||| |||||||| :::|||||::
Db         86 TLTALGAEVRWASCNIFSTQDEAAAAIVVGSGTVEEPAGVPVFAWKGESLEEYWWCINQI 145

Qy        216 VNMDGWQANMILDDGGDLTHWVYK---------------------------------- 239
               :      |||||||| |  | :
Db        146 FSWGDELPNMILDDGGDATMAVIRGREYEQAGLVPPAEANDSDEYIAFLGMLREVLAAEP 205

Qy        240 -KYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESIL 298
               |:  : : ::|: ||: |||||||  :: | |  |||||||:||| |||| |  | |::
Db        206 GKWGKIAEAVKGVTEETTTGVHRLYHPAEEGVLPFPAMNVNDAVTKSKFDNKYGTRHSLI 265

Qy        299 DGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVV 358
              ||: | ||:: ||| |:|||||:|||||  |    || | :|| ||| |||| |||: ||
Db        266 DGINRATDMLMGGKNVLVCGYGDVGKGCAEAFDGQGARVKVTEADPINALQALMDGYSVV 325

Qy        359 KLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTS-LRTPELTWE 417
               ::| |   |:||| ||||:::: | : :||:  :: |:|| : |||: | |   ::|
Db        326 TVDEAIEDADIVITATGNKDIISFEQMLKMKDHALLGNIGHFDNEIDMHSLLHRDDVTRT 385

Qy        418 RVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGR 476
               :: |||   :  |: :::|:|||||||  :|  |:||:| :   | :| |||:   ||:
Db        386 TIKPQVDEFTFSTGRSIIVLSEGRLLNLGNATGHPSFVMSNSFADQTIAQIELFQ-NEGQ 444

Qy        477 YKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              |: :|| ||| :|| || :|: :    ||||| :||:|:|::  ||||| :|||
Db        445 YENEVYRLPKVLDEKVARIHVEALGGQLTELTKEQAEYIGVDVAGPFKPEHYRY 498


RESULT 13
US-09-328-352-4600
; Sequence 4600, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4600
;   LENGTH: 463
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-4600

  Query Match             38.2%;  Score 1062;  DB 2;  Length 463;
  Best Local Similarity   47.6%;  Pred. No. 4.8e-111;
  Matches  216;  Conservative   84;  Mismatches  126;  Indels   28;  Gaps    6;

Qy        103 SNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLI 162
              ::: | :|  |::||:||::|| :| ||| ||||    ||||||||:|| |:| ||||||
Db         12 TDYKVADISLADYGRKEIKLAEAEMPALIGLRKRYAASKPLAGAKILGCIHMTIQTAVLI 71

Qy        163 ETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDG-- 220
              |||  |||: ||::|||:|||:  |||:|  |: |||||||:|::: ||::: :|::|
Db         72 ETLVELGAEVRWTSCNIFSTQDHAAAAIAARGIPVFAWKGETEEEYVWCLEQQINVNGQP 131

Qy        221 WQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVND 280
              | |||||||||||  |::||| : ::| || ||: ||| || :: | | | |||:||||
Db        132 WDANMILDDGGDLTALVHEKYPALLERIHGITEETTTGVQRLIEMWKDGTLKVPAINVND 191

Qy        281 SVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYIT 340
              |||| | || | || |: | :|| ||::  |:: :| |||:||||   :|:  | || :|
Db        192 SVTKSKNDNKYGCRHSLNDAIKRATDMLLSGRRALVIGYGDVGKGSAQSLRQEGMIVRVT 251

Qy        341 EIDPICALQACMDGFRVVK----------LNEVIRQVDVVITCTGNKNVVTREHLD 386
              |:|||||:||||||: ||              ::::   |:|:| ||| :|     ||
Db        252 EVDPICAMQACMDGYEVVSPYKNGVQTGKKEDINHDLLGNTDLVVTTTGNYHVCDAAMLD 311

Qy        387 RMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKR--VVLLAEGRLLN 444
               :|   :|||:|| :|||   ||      | |: ||  |   : :   ::||:||||:|
Db        312 SLKAGAVVCNIGHFDTEIDTAYLR--GYKWVEVKPQVHQVYRSEDENNYLILLSEGRLVN 369

Qy        445 LSCST-VPTFVLSITATTQALALIELYN-------APEGRYKQDVYLLPKKMDEYVASLH 496
              |  :|  |: |:  :   | |  | |:         | |   |  | :||||:|| ||:
Db        370 LGNATGHPSRVMDGSFANQVLGQIHLFQEKFADLPASEKAAKIRVEVLPKKLDEEVAAAM 429

Qy        497 LPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              :  |    ||:|| :|| |||:   |||| : |:|
Db        430 VAGFGGVLTQLTQEQADYLGIAVEGPFKSDAYKY 463


RESULT 14
US-09-759-990A-7
; Sequence 7, Application US/09759990A
; Patent No. 6713273
; GENERAL INFORMATION:
;  APPLICANT: AntiCancer, Inc.
;  APPLICANT:  Xu, Mingxu
;  APPLICANT:  Han, Oinghong
;  TITLE OF INVENTION: HIGH EXPRESSION AND PRODUCTION OF HIGN
;  TITLE OF INVENTION:  SPECIFIC ACTIVITY RECOMBINANT S-ADENOSYLHOMOCYSTAINASE
;  TITLE OF INVENTION:  (SAHH) AND IMPROVED ASSAYS FOR S-ADENOSYLMETHIONINE (SAM)
;  FILE REFERENCE: 31276-20026.00
;  CURRENT APPLICATION NUMBER: US/09/759,990A
;  CURRENT FILING DATE:  2001-01-12
;  PRIOR APPLICATION NUMBER: US 60/176,444
;  PRIOR FILING DATE: 2000-01-14
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
```

```
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-990A-7

  Query Match             37.1%;  Score 1032;  DB 2;  Length 486;
  Best Local Similarity   44.4%;  Pred. No. 1.3e-107;
  Matches  210;  Conservative   83;  Mismatches  130;  Indels   50;  Gaps    7;

Qy        107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
              : :|     ||:|: :||::|  |: ||:|    ||| | :| |  |:| |||||||||
Db         15 IADINLHVLGRKELTLAEKEMPGLMVLRERYSASKPLKGVRISGSLHMTVQTAVLIETLT 74

Qy        167 ALGAQCRWSACNIYSTQNEVAAALA---------EAGVAVFAWKGESEDDFWWCIDRCVN 217
              ||||  ||::|||:|||:  |||:           ||: |||||||:  ::|    | :
Db         75 ALGADVRWASCNIFSTQDTAAAAIVVGPTGTPEKPAGIPVFAWKGETLPEYWENTYRALT 134

Qy        218 M-DGWQANMILDDGGDLTHWVYKKY-------------------------PNVFKK--- 247
                ||     ::|||||  |  : | :                          || :
Db        135 WPDGQGPQQVVDDGGDATLLISKGFEFETAGAVPEPTEADNLEYRCVLATLKQVFNQDKN 194

Qy        248 --------IRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILD 299
                    : |: ||: |||||||| | |||  ||:||||:||| ||||:| || |::|
Db        195 HWHTVAAGMNGVSEETTTGVHRLYQLEKEGKLLFPAINVNDAVTKSKFDNIYGCRHSLID 254

Qy        300 GLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVK 359
              |: | :||| |||  :| |||:|||||  :|:  || | |||:|||||||| |:|::| :
Db        255 GINRASDVMIGGKTALVMGYGDVGKGCAQSLRGQGARVIITELDPICALQAAMEGYQVRR 314

Qy        360 LNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSL-RTPELTWER 418
              : ||:: ||: :||||| :::: : : :||:  || |:|| : |||    | : | :
Db        315 IEEVVKDVDIFVTCTGNCDIISVDMMAQMKDKAIVGNIGHFDNEIDTDGLMKYPGIKHIP 374

Qy        419 VRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGRY 477
              :: : |   :|||  ::|||||||||| |:|  |:||:|:: | | || ::||    |
Db        375 IKPEYDMWEFPDGHAILLLAEGRLLNLGCATGHPSFVMSMSFTNQTLAQLDLYE-KRGNL 433

Qy        478 KQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              :: || ||| :|| || ||| | | |||:||  || |: :   ||:| : |||
Db        434 EKKVYTLPKHLDEEVARLHLGSLDVHLTKLTQKQADYINVPVEGPYKSDAYRY 486


RESULT 15
US-09-489-039A-8393
; Sequence 8393, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 8393
;   LENGTH: 387
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8393

  Query Match             14.7%;  Score 409.5;  DB 2;  Length 387;
  Best Local Similarity   31.9%;  Pred. No. 4.2e-37;
  Matches  119;  Conservative   61;  Mismatches  170;  Indels   23;  Gaps    9;

Qy        118 REIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSAC 177
              :|:  | |:|   :   ::      |:| ::  | |:  :   |:: :   |||   : |
Db         27 KEVAWASQNMPRTL---RQVAALPDLSGVRLACCMHLDMKMIPLVQGILDKGAQVFLTTC 83

Qy        178 NIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWV 237
              |  : |::| | | | |    ||:  |: |:    :: :    ||   : : | |:|  :`
Db         84 NPTTVQDDVVAWLVERGAEACAWRNMSDADWQQSWEKAI---AWQPTHLCEMGADITTLL 140

Qy        238 YKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESI 297
              :::    |  |   :| : :||:||  :          |  | :|   |:   | :    :
Db        141 HQR--GEFGNIVAGLEATGSGVNRLGDIQPG----YPIFNWDDLPVKEGLHNRHMVGLTA 194

Qy        298 LDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRV 357
                   :|| :    |:|:| ||| ||:|  || || |  | : ||||   ||| ||: |
Db        195 WHTFFQTTHLTLHEKKVLVIGYGLVGQGVAAAAKAFGGQVMVAEIDPARRLQAAYDGWHV 254

Qy        358 VKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWE 417
              | | | |  ||| | || |||| |: ||| |    : |:||   |||   ||  :   |
Db        255 VDLQEAIASADVVATATGGKNVVNRQALDRAKAGVFILNVGHVAEEIDGDYLR--QYPQE 312

Qy        418 RVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIE---LYNAPE 474
              |   ::    | | | ||| | :|||:      |  |: |   ||::      :  :
Db        313 EVMPYINAYRMAD-KTVYLLANGSMLNLTAG----FGDSLNAFDVTLAVMASGIRHIVTD 367

Qy        475 G-RYKQDVYLLPK 486
              | |   |||||:
Db        368 GMRAPAKVYLLPQ 380


Search completed: December 14, 2005, 08:21:47
Job time : 38.6183 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:21:03 ; Search time 8.35962 Seconds
(without alignments)
426.936 Million cell updates/sec

Title: US-10-683-610-1
Perfect score: 2781
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....QAKYLGLNKNGPFKPNYYRY 530

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 51441 seqs, 6734002 residues

Total number of hits satisfying chosen parameters: 51441

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1151	41.4	432	6	US-10-836-953-1	Sequence 1, Appli
2	1074	38.6	478	7	US-11-055-822-220	Sequence 220, App
3	937	33.7	432	7	US-11-055-822-224	Sequence 224, App
4	107	3.8	35	7	US-11-055-822-222	Sequence 222, App
5	106	3.8	711	7	US-11-087-100-18	Sequence 18, Appl
6	106	3.8	711	7	US-11-087-084-18	Sequence 18, Appl
7	106	3.8	711	7	US-11-087-085-18	Sequence 18, Appl
8	106	3.8	2910	7	US-11-087-100-2	Sequence 2, Appli
9	106	3.8	2910	7	US-11-087-084-2	Sequence 2, Appli
10	106	3.8	2910	7	US-11-087-085-2	Sequence 2, Appli
11	97.5	3.5	1406	6	US-10-995-561-530	Sequence 530, App
12	97	3.5	1405	6	US-10-995-561-529	Sequence 529, App
13	96.5	3.5	1436	6	US-10-995-561-531	Sequence 531, App
14	95	3.4	414	7	US-11-089-551A-28	Sequence 28, Appl
15	90	3.2	560	7	US-11-034-569-6	Sequence 6, Appli
16	89.5	3.2	345	6	US-10-467-657-1374	Sequence 1374, Ap
17	88.5	3.2	564	7	US-11-034-569-17	Sequence 17, Appl
18	88	3.2	512	6	US-10-467-657-6780	Sequence 6780, Ap
19	87.5	3.1	5179	7	US-11-108-172-1068	Sequence 1068, Ap
20	86	3.1	366	6	US-10-821-234-1447	Sequence 1447, Ap
21	86	3.1	817	6	US-10-793-626-2948	Sequence 2948, Ap
22	85.5	3.1	1083	7	US-11-113-751-40	Sequence 40, Appl
23	84.5	3.0	1081	7	US-11-113-751-38	Sequence 38, Appl
24	84.5	3.0	1116	7	US-11-113-751-32	Sequence 32, Appl
25	84.5	3.0	1118	7	US-11-113-751-42	Sequence 42, Appl
26	84.5	3.0	1151	7	US-11-113-751-36	Sequence 36, Appl
27	84.5	3.0	1153	7	US-11-113-751-44	Sequence 44, Appl
28	84.5	3.0	1189	7	US-11-074-176-134	Sequence 134, App
29	84	3.0	410	6	US-10-858-730-233	Sequence 233, App
30	84	3.0	1116	7	US-11-113-751-34	Sequence 34, Appl
31	83.5	3.0	2897	6	US-10-499-715-2	Sequence 2, Appli
32	82.5	3.0	805	6	US-10-518-599-24	Sequence 24, Appl
33	82	2.9	421	6	US-10-763-712A-7	Sequence 7, Appli
34	81.5	2.9	446	7	US-11-196-475-150	Sequence 150, App
35	81	2.9	1044	7	US-11-091-668-2	Sequence 2, Appli
36	80.5	2.9	543	6	US-10-689-742-78	Sequence 78, Appl
37	80.5	2.9	1279	6	US-10-793-626-3188	Sequence 3188, Ap
38	80	2.9	244	7	US-11-054-515-1933	Sequence 1933, Ap
39	80	2.9	1452	6	US-10-995-561-778	Sequence 778, App
40	80	2.9	3803	6	US-10-995-561-773	Sequence 773, App
41	80	2.9	3960	6	US-10-995-561-771	Sequence 771, App
42	80	2.9	5335	6	US-10-995-561-777	Sequence 777, App
43	80	2.9	5406	6	US-10-995-561-774	Sequence 774, App
44	80	2.9	5415	6	US-10-995-561-779	Sequence 779, App
45	80	2.9	5464	6	US-10-995-561-775	Sequence 775, App

ALIGNMENTS

RESULT 1
US-10-836-953-1
; Sequence 1, Application US/10836953
; Publication No. US20050244912A1
; GENERAL INFORMATION:
; APPLICANT: YUAN, Chong-Sheng
; APPLICANT: YE, Qi-Zhuang
; APPLICANT: XIE, Sheng,Xue
; TITLE OF INVENTION: METHODS FOR ASSAYING INHIBITORS OF
; TITLE OF INVENTION: S-ADENOSYLHOMOCYSTEINE (SAH) HYDROLASE AND
; TITLE OF INVENTION: S-ADENOSYLMETHIONINE (SAM)-DEPENDENT METHYLTRANSFERASE
; FILE REFERENCE: 466992001600
; CURRENT APPLICATION NUMBER: US/10/836,953
; CURRENT FILING DATE: 2004-04-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-836-953-1

Query Match 41.4%; Score 1151; DB 6; Length 432;
Best Local Similarity 51.3%; Pred. No. 4.9e-93;
Matches 218; Conservative 85; Mismatches 120; Indels 2; Gaps 2;

```
Qy      107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
            | :|  | :||: ::||| :|  |: :|:|    ||| ||:| || |:| :||||||||
Db        9 VADIGLAAWGRKALDIAENEMPGLMRMRERYSASKPLKGARIAGCLHMTVETAVLIETLV 68

Qy      167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMI 226
            |||: :||:|||:|||| |||:|:||: |:|||||:::: |||:: :       |||
Db       69 TLGAEVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLWCIEQTLYFKDGPLNMI 128

Qy      227 LDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQK 286
            ||||||||: :: ||| :   |||| ||: |||| ||::   | | |||:|||||||| |
Db      129 LDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVHNLYKMMANGILKVPAINVNDSVTKSK 188

Qy      287 FDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPIC 346
            ||||| |||||::||:|| |||| ||  || |||:|||||  ||:  || | |||||||
Db      189 FDNLYGCRESLIDGIKRATDVMIAGKVAVVAGYGDVGKGCAQALRGFGARVIITEIDPIN 248

Qy      347 ALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDV 406
            |||| |:|: |  ::|  :: :: :| ||  :::   | ::||:  ||||:|| : ||||
Db      249 ALQAAMEGYEVTTMDEACQEGNIFVTTTGCIDIILGRHFEQMKDDAIVCNIGHFDVEIDV 308

Qy      407 TSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALA 465
              |   :   :: |||     :|:|::|||||||:|| |:  |:||:| : | | :|
```

```
Db        309 KWLNENAVEKVNIKPQVDRYRLKNGRRIILLAEGRLVNLGCAMGHPSFVMSNSFTNQVMA 368

Qy        466 LIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKP 525

Db        369 QIELWTHPD-KYPVGVHFLPKKLDEAVAEAHLGKLNVKLTKLTEKQAQYLGMSCDGPFKP 427

Qy        526 NYYRY 530

Db        428 DHYRY 432


RESULT 2
US-11-055-822-220
; Sequence 220, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
;  APPLICANT: Pompejus, Markus
;  APPLICANT:  Kroger, Burkhard
;  APPLICANT:  Schroder, Hartwig
;  APPLICANT:  Zelder, Oskar
;  APPLICANT:  Haberhauer, Gregor
;  TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
;  TITLE OF INVENTION:  METABOLIC PATHWAY PROTEINS
;  FILE REFERENCE: BGI-121CPCN
;  CURRENT APPLICATION NUMBER: US/11/055,822
;  CURRENT FILING DATE:  2005-02-11
;  PRIOR APPLICATION NUMBER: 09/606,740
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: 60/141,031
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: 60/142,101
;  PRIOR FILING DATE: 1999-07-02
;  PRIOR APPLICATION NUMBER: 60/148,613
;  PRIOR FILING DATE: 1999-08-12
;  PRIOR APPLICATION NUMBER: 60/187,970
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: DE 19930476.9
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: DE 19931415.2
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931418.7
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931419.5
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931420.9
;  PRIOR FILING DATE: 1999-07-08
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 220
;   LENGTH: 478
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-11-055-822-220

  Query Match             38.6%;  Score 1074;  DB 7;  Length 478;
  Best Local Similarity   45.4%;  Pred. No. 3e-86;
  Matches  215;  Conservative   86;  Mismatches  125;  Indels   48;  Gaps    5;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163

Db          6 DFKVADLSLAEAGRHQIRLAEYEMPGLMQLRKEFADEQPLKGARIAGSIHMTVQTAVLIE 65

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALA--------EAGVAVFAWKGESEDDFWWCIDRC 215

Db         66 TLTALGAEVRWASCNIFSTQDEAAAAIVVGSGTVEEPAGVPVFAWKGESLEEYWWCINQI 125

Qy        216 VNMDGWQANMILDDGGDLTHWVYK---------------------------------- 239

Db        126 FSWGDELPNMILDDGGDATMAVIRGREYEQAGLVPPAEANDSDEYIAFLGMLREVLAAEP 185

Qy        240 -KYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESIL 298

Db        186 GKWGKIAEAVKGVTEETTTGVHRLYHFAEEGVLPFPAMNVNDAVTKSKFDNKYGTRHSLI 245

Qy        299 DGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVV 358

Db        246 DGINRATDMLMGGKNVLVCGYGDVGKGCAEAFDGQGARVKVTEADPINALQALMDGYSVV 305

Qy        359 KLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTS-LRTPELTWE 417

Db        306 TVDEAIEDADIVITATGNKDIISFEQMLKMKDHALLGNIGHFDNEIDMHSLLHRDDVTRT 365

Qy        418 RVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGR 476

Db        366 TIKPQVDEFTFSTGRSIIVLSEGRLLNLGNATGHPSFVMSNSFADQTIAQIELFQ-NEGQ 424

Qy        477 YKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530

Db        425 YENEVYRLPKVLDEKVARIHVEALGGQLTELTKEQAEYIGVDVAGPFKPEHYRY 478


RESULT 3
US-11-055-822-224
; Sequence 224, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
;  APPLICANT: Pompejus, Markus
;  APPLICANT:  Kroger, Burkhard
;  APPLICANT:  Schroder, Hartwig
;  APPLICANT:  Zelder, Oskar
;  APPLICANT:  Haberhauer, Gregor
;  TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
;  TITLE OF INVENTION:  METABOLIC PATHWAY PROTEINS
;  FILE REFERENCE: BGI-121CPCN
;  CURRENT APPLICATION NUMBER: US/11/055,822
;  CURRENT FILING DATE:  2005-02-11
;  PRIOR APPLICATION NUMBER: 09/606,740
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: 60/141,031
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: 60/142,101
;  PRIOR FILING DATE: 1999-07-02
;  PRIOR APPLICATION NUMBER: 60/148,613
;  PRIOR FILING DATE: 1999-08-12
;  PRIOR APPLICATION NUMBER: 60/187,970
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: DE 19930476.9
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: DE 19931415.2
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931418.7
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931419.5
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931420.9
;  PRIOR FILING DATE: 1999-07-08
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 224
;   LENGTH: 432
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-11-055-822-224

  Query Match             33.7%;  Score 937;  DB 7;  Length 432;
  Best Local Similarity   44.6%;  Pred. No. 2.3e-74;
  Matches  191;  Conservative   76;  Mismatches  113;  Indels   48;  Gaps    5;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163

Db          6 DFKVADLSLAEAGRHQIRLAEYEMPGLMQLRKEFADEQPLKGARIAGSIHMTVQTAVLIE 65

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALA--------EAGVAVFAWKGESEDDFWWCIDRC 215

Db         66 TLTALGAEVRWASCNIFSTQDEAAAAIVVGSGTVEEPAGVPVFAWKGESLEEYWWCINQI 125
```

```
Qy        216 VNMDGWQANMILDDGGDLTHWVYK---------------------------------- 239
              :          |||||||||  |   | :
Db        126 FSWGDELPNMILDDGGDATMAVIRGREYEQAGLVPPAEANDSDEYIAFLGMLREVLAAEP 185

Qy        240 -KYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESIL 298
               |:  : : ::|: ||: |||||||  :: | |  |||||||:||| |||| |  | |::
Db        186 GKWGKIAEAVKGVTEETTTGVHRLYHFAEEGVLPFPAMNVNDAVTKSKFDNKYGTRHSLI 245

Qy        299 DGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVV 358
               ||: | ||:: ||| |:|||||:|||||  |     || | :|| ||| |||| |||: ||
Db        246 DGINRATDMLMGGKNVLVCGYGDVGKGCAEAFDGQGARVKVTEADPINALQALMDGYSVV 305

Qy        359 KLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTS-LRTPELTWE 417
               ::| |   |:||| |||||::: | : :||:  :: |:|| : |||: | |   ::|
Db        306 TVDEAIEDADIVITATGNKDIISFEQMLKMKDHALLGNIGHFDNEIDMHSLLHRDDVTRT 365

Qy        418 RVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGR 476
               :: |||   :  |: :::|:|||||||  :|  |:||:| :   | :| |||:   ||:
Db        366 TIKPQVDEFTFSTGRSIIVLSEGRLLNLGNATGHPSFVMSNSFADQTIAQIELFQ-NEGQ 424

Qy        477 YKQDVYLL 484
              |: :|| |
Db        425 YENEVYRL 432
```

RESULT 4
US-11-055-822-222
; Sequence 222, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 222
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-222

Query Match 3.8%; Score 107; DB 7; Length 35;
Best Local Similarity 51.4%; Pred. No. 0.0016;
Matches 18; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

```
Qy        496 HLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
```

```
              |: :    ||||| :||:|:|::  ||||| :|||
Db          1 HVEALGGQLTELTKEQAEYIGVDVAGPFKPEHYRY 35
```

RESULT 5
US-11-087-100-18
; Sequence 18, Application US/11087100
; Publication No. US20050266440A1
; GENERAL INFORMATION:
; APPLICANT: Metz, James
; APPLICANT: Barclay, William
; APPLICANT: Flatt, James
; APPLICANT: Kuner, Jerry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Syntha:
; TITLE OF INVENTION: System and Uses Thereof
; FILE REFERENCE: 2997-29
; CURRENT APPLICATION NUMBER: US/11/087,100
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: 09/231,899
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/284,066
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/298,796
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/323,269
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Schizochytrium sp.
US-11-087-100-18

Query Match 3.8%; Score 106; DB 7; Length 711;
Best Local Similarity 22.1%; Pred. No. 0.15;
Matches 107; Conservative 60; Mismatches 157; Indels 160; Gaps 25;

```
Qy         99 SKGSSNFCVKNI-------KQAEFGRREIEIAEQDMSALISLRKRAQ------------ 138
              ::| :  |:: |       |    || ::  :|    | |:  | |
Db        162 ARGITPLCIREITRQIAGGKYILLGRSKVSASEPAWCAGITDEKAVQKAATQELKRAFSA 221

Qy        139 --GEKPL--AGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAG 194
                 | ||   |  |:||      :    |  : ||| :  :|:|:: |   :|| |: :|
Db        222 GEGPKPTPRAVTKLVGSVLGAREVRSSIAAIEALGGKAIYSSCDVNSAA-DVAKAVRDA- 279

Qy        195 VAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEE 254
                       |      :   |:  |    :| |       : || |: |  : |
Db        280 ----------ESQLGARVSGIVHASG-----VLRD-----RLIEKKLPDEFDAVFG---T 316

Qy        255 SVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQV 314
               |||:  |           |   |:   |       |:    | | |
Db        317 KVTGLENLL-------AAVDRANLKHMV-------LF----SSLAGF------------- 345

Qy        315 VVCGYGEVGKGCCA----ALKALG-AIVYITEIDPICALQACMDGFRVV-KLNEVIRQVD 368
               :| ||:   |    ||  :|  :    :  ||      ||  |  :| :  :::
Db        346 ----HGNVGQSDYAMANEALNKMGLELAKDVSVKSICF--GPWDGGMVTPQLKKQFQEMG 399

Qy        369 V-VITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPE-------LTWERVR 420
               | :|   |  : | |          ||  :| |  || | : |||       :|  |
Db        400 VQIIPREGGADTVAR----------IV--LGSSPAEILVGNWRTPSKKVGSDTITLHRKI 447

Qy        421 SQ------VDHVIWPDGKRV--VLLAEGRL---------------------------- 442
               |       ||||   |:||  : || | |
Db        448 SAKSNPFLEDHVI--QGRRVLPMTLAIGSLAETCLGLFPGYSLWAIDDAQLFKGVTVDGD 505

Qy        443 ----LNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLP 498
                   : |: || |:  ::: || :  :  :|  |     |:  : |  :       |::  |
Db        506 VNCEVTLTPSTAPSGRVNVQATLKTFSSGKLVPA----YRAVIVLSNQGAPPANATMQPP 561

Qy        499 SFDA 502
               | ||
Db        562 SLDA 565
```

```
RESULT 6
US-11-087-084-18
; Sequence 18, Application US/11087084
; Publication No. US20050273883A1
; GENERAL INFORMATION:
;  APPLICANT: Metz, James
;  APPLICANT:  Barclay, William
;  APPLICANT:  Flatt, James
;  APPLICANT:  Kuner, Jerry
;  TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase
;  TITLE OF INVENTION:  System and Uses Thereof
;  FILE REFERENCE: 2997-29
;  CURRENT APPLICATION NUMBER: US/11/087,084
;  CURRENT FILING DATE:  2005-03-21
;  PRIOR APPLICATION NUMBER: 09/231,899
;  PRIOR FILING DATE: 1999-01-14
;  PRIOR APPLICATION NUMBER: 60/284,066
;  PRIOR FILING DATE: 2001-04-16
;  PRIOR APPLICATION NUMBER: 60/298,796
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/323,269
;  PRIOR FILING DATE: 2001-09-18
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
;   LENGTH: 711
;   TYPE: PRT
;   ORGANISM: Schizochytrium sp.
US-11-087-084-18

  Query Match             3.8%;  Score 106;  DB 7;  Length 711;
  Best Local Similarity   22.1%;  Pred. No. 0.15;
  Matches  107;  Conservative   60;  Mismatches  157;  Indels  160;  Gaps   25;

Qy         99 SKGSSNFCVKNI-------KQAEFGRREIEIAEQDMSALISLRKRAQ------------ 138
              ::| :  |:: |       |    || ::  :|    | |:  |  |
Db        162 ARGITPLCIREITRQIAGGKYILLGRSKVSASEPAWCAGITDEKAVQKAATQELKRAFSA 221

Qy        139 --GEKPL--AGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAG 194
                | ||   |  |:||      :    |  : ||| :  :|:|:: |   :|| |: :|
Db        222 GEGPKPTPRAVTKLVGSVLGAREVRSSIAAIEALGGKAIYSSCDVNSAA-DVAKAVRDA- 279

Qy        195 VAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEE 254
                       |      :   |:  |     :| |       : || |: |  : |
Db        280 ----------ESQLGARVSGIVHASG-----VLRD-----RLIEKKLPDEFDAVFG---T 316

Qy        255 SVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQV 314
               |||:  |          |   |:   |       |:    | | |
Db        317 KVTGLENLL-------AAVDRANLKHMV-------LF----SSLAGF------------- 345

Qy        315 VVCGYGEVGKGCCA----ALKALG-AIVYITEIDPICALQACMDGFRVV-KLNEVIRQVD 368
                :| ||:   |    ||  :|  :     :  ||       ||  |  :| :  :::
Db        346 ----HGNVGQSDYAMANEALNKMGLELAKDVSVKSICF--GPWDGGMVTPQLKKQFQEMG 399

Qy        369 V-VITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPE-------LTWERVR 420
              | :|   |  : | |           ||  :| |  || | : |||        :|  |
Db        400 VQIIPREGGADTVAR----------IV--LGSSPAEILVGNWRTPSKKVGSDTITLHRKI 447

Qy        421 SQ------VDHVIWPDGKRV--VLLAEGRL---------------------------- 442
              |       ||||   |:||  : || | |
Db        448 SAKSNPFLEDHVI--QGRRVLPMTLAIGSLAETCLGLFPGYSLWAIDDAQLFKGVTVDGD 505

Qy        443 ----LNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLP 498
                  : |: || |:  ::: || :  :  :| |    |:  : |  :       |:: |
Db        506 VNCEVTLTPSTAPSGRVNVQATLKTFSSGKLVPA----YRAVIVLSNQGAPPANATMQPP 561

Qy        499 SFDA 502
              | ||
Db        562 SLDA 565
```

```
RESULT 7
US-11-087-085-18
; Sequence 18, Application US/11087085
; Publication No. US20050273884A1
; GENERAL INFORMATION:
;  APPLICANT: Metz, James
;  APPLICANT:  Barclay, William
;  APPLICANT:  Flatt, James
;  APPLICANT:  Kuner, Jerry
;  TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthas
;  TITLE OF INVENTION:  System and Uses Thereof
;  FILE REFERENCE: 2997-29
;  CURRENT APPLICATION NUMBER: US/11/087,085
;  CURRENT FILING DATE:  2005-03-21
;  PRIOR APPLICATION NUMBER: 09/231,899
;  PRIOR FILING DATE: 1999-01-14
;  PRIOR APPLICATION NUMBER: 60/284,066
;  PRIOR FILING DATE: 2001-04-16
;  PRIOR APPLICATION NUMBER: 60/298,796
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/323,269
;  PRIOR FILING DATE: 2001-09-18
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
;   LENGTH: 711
;   TYPE: PRT
;   ORGANISM: Schizochytrium sp.
US-11-087-085-18

  Query Match             3.8%;  Score 106;  DB 7;  Length 711;
  Best Local Similarity   22.1%;  Pred. No. 0.15;
  Matches  107;  Conservative   60;  Mismatches  157;  Indels  160;  Gaps   25;

Qy         99 SKGSSNFCVKNI-------KQAEFGRREIEIAEQDMSALISLRKRAQ------------ 138
              ::| :  |:: |       |    || ::  :|    | |:  |  |
Db        162 ARGITPLCIREITRQIAGGKYILLGRSKVSASEPAWCAGITDEKAVQKAATQELKRAFSA 221

Qy        139 --GEKPL--AGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAG 194
                | ||   |  |:||      :    |  : ||| :  :|:|:: |   :|| |: :|
Db        222 GEGPKPTPRAVTKLVGSVLGAREVRSSIAAIEALGGKAIYSSCDVNSAA-DVAKAVRDA- 279

Qy        195 VAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEE 254
                       |      :   |:  |     :| |       : || |: |  : |
Db        280 ----------ESQLGARVSGIVHASG-----VLRD-----RLIEKKLPDEFDAVFG---T 316

Qy        255 SVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQV 314
               |||:  |          |   |:   |       |:    | | |
Db        317 KVTGLENLL-------AAVDRANLKHMV-------LF----SSLAGF------------- 345

Qy        315 VVCGYGEVGKGCCA----ALKALG-AIVYITEIDPICALQACMDGFRVV-KLNEVIRQVD 368
                :| ||:   |    ||  :|  :     :  ||       ||  |  :| :  :::
Db        346 ----HGNVGQSDYAMANEALNKMGLELAKDVSVKSICF--GPWDGGMVTPQLKKQFQEMG 399

Qy        369 V-VITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPE-------LTWERVR 420
              | :|   |  : | |           ||  :| |  || | : |||        :|  |
Db        400 VQIIPREGGADTVAR----------IV--LGSSPAEILVGNWRTPSKKVGSDTITLHRKI 447

Qy        421 SQ------VDHVIWPDGKRV--VLLAEGRL---------------------------- 442
              |       ||||   |:||  : || | |
Db        448 SAKSNPFLEDHVI--QGRRVLPMTLAIGSLAETCLGLFPGYSLWAIDDAQLFKGVTVDGD 505

Qy        443 ----LNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLP 498
                  : |: || |:  ::: || :  :  :| |    |:  : |  :       |:: |
Db        506 VNCEVTLTPSTAPSGRVNVQATLKTFSSGKLVPA----YRAVIVLSNQGAPPANATMQPP 561

Qy        499 SFDA 502
              | ||
Db        562 SLDA 565
```

```
RESULT 8
US-11-087-100-2
```

```
; Sequence 2, Application US/11087100
; Publication No. US20050266440A1
; GENERAL INFORMATION:
;  APPLICANT: Metz, James
;  APPLICANT:  Barclay, William
;  APPLICANT:  Flatt, James
;  APPLICANT:  Kuner, Jerry
;  TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase
;  TITLE OF INVENTION:  System and Uses Thereof
;  FILE REFERENCE: 2997-29
;  CURRENT APPLICATION NUMBER: US/11/087,100
;  CURRENT FILING DATE:  2005-03-21
;  PRIOR APPLICATION NUMBER: 09/231,899
;  PRIOR FILING DATE: 1999-01-14
;  PRIOR APPLICATION NUMBER: 60/284,066
;  PRIOR FILING DATE: 2001-04-16
;  PRIOR APPLICATION NUMBER: 60/298,796
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/323,269
;  PRIOR FILING DATE: 2001-09-18
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
;   LENGTH: 2910
;   TYPE: PRT
;   ORGANISM: Schizochytrium sp.
US-11-087-100-2

  Query Match             3.8%;  Score 106;  DB 7;  Length 2910;
  Best Local Similarity   22.1%;  Pred. No. 1.2;
  Matches  107;  Conservative   60;  Mismatches  157;  Indels  160;  Gaps   25;

Qy         99 SKGSSNFCVKNI-------KQAEFGRREIEIAEQDMSALISLRKRAQ------------ 138
              ::| :  |:: |       |  || ::  :|    | |:  | |
Db       2361 ARGITPLCIREITRQIAGGKYILLGRSKVSASEPAWCAGITDEKAVQKAATQELKRAFSA 2420

Qy        139 --GEKPL--AGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAG 194
               | ||   |  |:||      :    |  : ||| :  :|:|:: |   :|| |: :|
Db       2421 GEGPKPTPRAVTKLVGSVLGAREVRSSIAAIEALGGKAIYSSCDVNSAA-DVAKAVRDA- 2478

Qy        195 VAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEE 254
                       |      :   |:  |     :| |       : || |: |  : |
Db       2479 ----------ESQLGARVSGIVHASG-----VLRD-----RLIEKKLPDEFDAVFG---T 2515

Qy        255 SVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQV 314
              |||:  |          |   |:  |       |:    | | |
Db       2516 KVTGLENLL-------AAVDRANLKHMV-------LF----SSLAGF------------- 2544

Qy        315 VVCGYGEVGKGCCA----ALKALG-AIVYITEIDPICALQACMDGFRVV-KLNEVIRQVD 368
                :| ||:   |    ||  :|  :     :  ||      ||  |  :| :  :::
Db       2545 ----HGNVGQSDYAMANEALNKMGLELAKDVSVKSICF--GPWDGGMVTPQLKKQFQEMG 2598

Qy        369 V-VITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPE-------LTWERVR 420
               | :|   |  : | |          ||  :| |  || | : |||         :|  |
Db       2599 VQIIPREGGADTVAR----------IV--LGSSPAEILVGNWRTPSKKVGSDTITLHRKI 2646

Qy        421 SQ------VDHVIWPDGKRV--VLLAEGRL---------------------------- 442
               |       ||||   |:||  : || | |
Db       2647 SAKSNPFLEDHVI--QGRRVLPMTLAIGSLAETCLGLFPGYSLWAIDDAQLFKGVTVDGD 2704

Qy        443 ----LNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLP 498
                   : |: || |:  ::: || :  :  :|  |    |:  : |  :      |:: |
Db       2705 VNCEVTLTPSTAPSGRVNVQATLKTFSSGKLVPA----YRAVIVLSNQGAPPANATMQPP 2760

Qy        499 SFDA 502
               | ||
Db       2761 SLDA 2764


RESULT 9
US-11-087-084-2
; Sequence 2, Application US/11087084
; Publication No. US20050273883A1
; GENERAL INFORMATION:
;  APPLICANT: Metz, James
;  APPLICANT:  Barclay, William
;  APPLICANT:  Flatt, James
;  APPLICANT:  Kuner, Jerry
;  TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Syntha:
;  TITLE OF INVENTION:  System and Uses Thereof
;  FILE REFERENCE: 2997-29
;  CURRENT APPLICATION NUMBER: US/11/087,084
;  CURRENT FILING DATE:  2005-03-21
;  PRIOR APPLICATION NUMBER: 09/231,899
;  PRIOR FILING DATE: 1999-01-14
;  PRIOR APPLICATION NUMBER: 60/284,066
;  PRIOR FILING DATE: 2001-04-16
;  PRIOR APPLICATION NUMBER: 60/298,796
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/323,269
;  PRIOR FILING DATE: 2001-09-18
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
;   LENGTH: 2910
;   TYPE: PRT
;   ORGANISM: Schizochytrium sp.
US-11-087-084-2

  Query Match             3.8%;  Score 106;  DB 7;  Length 2910;
  Best Local Similarity   22.1%;  Pred. No. 1.2;
  Matches  107;  Conservative   60;  Mismatches  157;  Indels  160;  Gaps   25;

Qy         99 SKGSSNFCVKNI-------KQAEFGRREIEIAEQDMSALISLRKRAQ------------ 138
              ::| :  |:: |       |  || ::  :|    | |:  | |
Db       2361 ARGITPLCIREITRQIAGGKYILLGRSKVSASEPAWCAGITDEKAVQKAATQELKRAFSA 2420

Qy        139 --GEKPL--AGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAG 194
               | ||   |  |:||      :    |  : ||| :  :|:|:: |   :|| |: :|
Db       2421 GEGPKPTPRAVTKLVGSVLGAREVRSSIAAIEALGGKAIYSSCDVNSAA-DVAKAVRDA- 2478

Qy        195 VAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEE 254
                       |      :   |:  |     :| |       : || |: |  : |
Db       2479 ----------ESQLGARVSGIVHASG-----VLRD-----RLIEKKLPDEFDAVFG---T 2515

Qy        255 SVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQV 314
              |||:  |          |   |:  |       |:    | | |
Db       2516 KVTGLENLL-------AAVDRANLKHMV-------LF----SSLAGF------------- 2544

Qy        315 VVCGYGEVGKGCCA----ALKALG-AIVYITEIDPICALQACMDGFRVV-KLNEVIRQVD 368
                :| ||:   |    ||  :|  :     :  ||      ||  |  :| :  :::
Db       2545 ----HGNVGQSDYAMANEALNKMGLELAKDVSVKSICF--GPWDGGMVTPQLKKQFQEMG 2598

Qy        369 V-VITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPE-------LTWERVR 420
               | :|   |  : | |          ||  :| |  || | : |||         :|  |
Db       2599 VQIIPREGGADTVAR----------IV--LGSSPAEILVGNWRTPSKKVGSDTITLHRKI 2646

Qy        421 SQ------VDHVIWPDGKRV--VLLAEGRL---------------------------- 442
               |       ||||   |:||  : || | |
Db       2647 SAKSNPFLEDHVI--QGRRVLPMTLAIGSLAETCLGLFPGYSLWAIDDAQLFKGVTVDGD 2704

Qy        443 ----LNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLP 498
                   : |: || |:  ::: || :  :  :|  |    |:  : |  :      |:: |
Db       2705 VNCEVTLTPSTAPSGRVNVQATLKTFSSGKLVPA----YRAVIVLSNQGAPPANATMQPP 2760

Qy        499 SFDA 502
               | ||
Db       2761 SLDA 2764


RESULT 10
US-11-087-085-2
; Sequence 2, Application US/11087085
; Publication No. US20050273884A1
; GENERAL INFORMATION:
;  APPLICANT: Metz, James
```

```
;  APPLICANT:  Barclay, William
;  APPLICANT:  Flatt, James
;  APPLICANT:  Kuner, Jerry
;  TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase
;  TITLE OF INVENTION:  System and Uses Thereof
;  FILE REFERENCE: 2997-29
;  CURRENT APPLICATION NUMBER: US/11/087,085
;  CURRENT FILING DATE:  2005-03-21
;  PRIOR APPLICATION NUMBER: 09/231,899
;  PRIOR FILING DATE: 1999-01-14
;  PRIOR APPLICATION NUMBER: 60/284,066
;  PRIOR FILING DATE: 2001-04-16
;  PRIOR APPLICATION NUMBER: 60/298,796
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/323,269
;  PRIOR FILING DATE: 2001-09-18
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
;   LENGTH: 2910
;   TYPE: PRT
;   ORGANISM: Schizochytrium sp.
US-11-087-085-2

  Query Match             3.8%;  Score 106;  DB 7;  Length 2910;
  Best Local Similarity   22.1%;  Pred. No. 1.2;
  Matches  107;  Conservative   60;  Mismatches  157;  Indels  160;  Gaps   25;

Qy         99 SKGSSNFCVKNI-------KQAEFGRREIEIAEQDMSALISLRKRAQ------------ 138
              ::| :  |:: |       |    || ::  :|    | |:  |  |
Db       2361 ARGITPLCIREITRQIAGGKYILLGRSKVSASEPAWCAGITDEKAVQKAATQELKRAFSA 2420

Qy        139 --GEKPL--AGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAG 194
                | ||   |  |:||      :    |  : ||| :  :|:|:: |   :|| |: :|
Db       2421 GEGPKPTPRAVTKLVGSVLGAREVRSSIAAIEALGGKAIYSSCDVNSAA-DVAKAVRDA- 2478

Qy        195 VAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEE 254
                       |    :  |:  |    :| |      : || |: |  : |
Db       2479 ----------ESQLGARVSGIVHASG-----VLRD-----RLIEKKLPDEFDAVFG---T 2515

Qy        255 SVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQV 314
              |||:  |          |   |:   |       |:    | | |
Db       2516 KVTGLENLL-------AAVDRANLKHMV-------LF----SSLAGF------------ 2544

Qy        315 VVCGYGEVGKGCCA----ALKALG-AIVYITEIDPICALQACMDGFRVV-KLNEVIRQVD 368
                 :| ||:   |    ||  :|  :     :  ||      ||  |  :| :  :::
Db       2545 ----HGNVGQSDYAMANEALNKMGLELAKDVSVKSICF--GPWDGGMVTPQLKKQFQEMG 2598

Qy        369 V-VITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPE-------LTWERVR 420
              | :|   |  : | |          ||  :| |  || | : |||         :|  |
Db       2599 VQIIPREGGADTVAR----------IV--LGSSPAEILVGNWRTPSKKVGSDTITLHRKI 2646

Qy        421 SQ------VDHVIWPDGKRV--VLLAEGRL------------------------------ 442
              |      ||||   |:||  : || | |
Db       2647 SAKSNPFLEDHVI--QGRRVLPMTLAIGSLAETCLGLFPGYSLWAIDDAQLFKGVTVDGD 2704

Qy        443 ----LNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLP 498
                  : |: || |:  ::: || :  :  :|  |    |:  : |  :      |::  |
Db       2705 VNCEVTLTPSTAPSGRVNVQATLKTFSSGKLVPA----YRAVIVLSNQGAPPANATMQPP 2760

Qy        499 SFDA 502
              | ||
Db       2761 SLDA 2764


RESULT 11
US-10-995-561-530
; Sequence 530, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 530
;   LENGTH: 1406
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-995-561-530

  Query Match             3.5%;  Score 97.5;  DB 6;  Length 1406;
  Best Local Similarity   21.0%;  Pred. No. 2.3;
  Matches  115;  Conservative   68;  Mismatches  183;  Indels  181;  Gaps   34;

Qy         36 KAPKKQIQFADDMQEFTKFPTKTGRRSLS---------RSISQSSTDSYSSAASYTDSSD 86
              |  :|:  | | :::  |      |  |           : : | | : | |  :|  |::
Db       366 KVERKEDGFEDGVED-NKLKENMERACLMSLDITEHELQILEQQSQEEYLSDIAY-KSTE 423

Qy         87 DEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGA 146
                 :|| : :   |      : ::: :  |             | | ||      |
Db       424 KHLSPNDNENDTS-----YVIESDEDLE-----------MEMLKSLENLNSG------- 459

Qy        147 KIVGCTHITAQTAVLIETLCALGAQCRWSACNI-YSTQNEVAAALAEAGVAVFAWKGESE 205
                | ||                ::|      |:  |: |      ||        :|| :
Db       460 -TVEPTH---------------SKCLKMERNLGLPTKEEEEDDENEAN------EGEED 496

Qy        206 D--DFWWCIDR-----CVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEE---- 254
              |  || |          |: |              | :|  |  :| |  ::||
Db       497 DDKDFLWPAPNEEQVTCLKM-------------YFGHSSFK--PVQWKVIHSVLEERRDN 541

Qy        255 ---SVTGVHR--LYQ-----LSKAGKLCVPAMN-VNDSVTKQKFDNLYCC------RESI 297
                   ||  :   :|      : | | :  | :: : | | :  |  |:  |       |::
Db       542 VAVMATGYGKSLCFQYPPVYVGKIGLVISPLISLMEDQVLQLKMSNIPACFLGSAQSENV 601

Qy        298 LDGLKRTTDVMPGGKQVVVCGYGEVGKGCCAALKALGAIVYITEI---DPICALQ---AC 351
              |       ||:   ||  :|     |    |     |: | | : || |    :  | :
Db       602 L------TDIKL-GKYRIVYVTPEYCSGNMGLLQQLEADIGITLIAVDEAHCISEWGHDF 654

Qy        352 MDGFRVV-KLNEVIRQVDVV-ITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSL 409
              | || :  |  :  | :| :| | : ::  || : |  |           | :|  |
Db       655 RDSFRKLGSLKTALPMVPIVALTATASSSI--REDIVRCLNL--------RNPQITCTGF 704

Qy        410 RTPELTWERVRSQVDHVI-------------WP-DGKRVV----------LLAEGRLLNL 445
              | |  | || :  :::               | :|  ::          :  | | |||
Db       705 DRPNLYLE-VRRKTGNILQDLQPFLVKTSSHWEFEGPTIIYCPSRKMTQQVTGELRKLNL 763

Qy        446 SCSTVPT-------------FV---LSITATTQALALIELYNAPEGRYKQDV-----YLL 484
              || |                ||   :    | |  :          |  | |:     |
Db       764 SCGTYHAGMSFSTRKDIHHRFVRDEIQCVIATIAFGM--------GINKADIRQVIHYGA 815

Qy        485 PKKMDEY 491
              || |: |
Db       816 PKDMESY 822


RESULT 12
US-10-995-561-529
; Sequence 529, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 529
;   LENGTH: 1405
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-995-561-529

  Query Match             3.5%;  Score 97;  DB 6;  Length 1405;
  Best Local Similarity   21.0%;  Pred. No. 2.5;
  Matches  115;  Conservative   67;  Mismatches  183;  Indels  182;  Gaps   34;

Qy         36 KAPKKQIQFADDMQEFTKFPTKTGRRSLS---------RSISQSSTDSYSSAASYTDSSD 86
Db        366 KVERKEDGFEDGVED-NKLKENMERACLMSLDITEHELQILEQQSQEEYLSDIAY--KST 422

Qy         87 DEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGA 146
Db        423 EHLSPNDNENDTS-----YVIESDEDLE-----------MEMLKSLENLNSG------- 458

Qy        147 KIVGCTHITAQTAVLIETLCALGAQCRWSACNI-YSTQNEVAAALAEAGVAVFAWKGESE 205
Db        459 -TVEPTH---------------SKCLKMERNLGLPTKEEEEDDENEAN------EGEED 495

Qy        206 D--DFWWCIDR-----CVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEE---- 254
Db        496 DDKDFLWPAPNEEQVTCLKM------------YFGHSSFK--PVQWKVIHSVLEERRDN 540

Qy        255 ---SVTGVHR--LYQ-----LSKAGKLCVPAMN-VNDSVTKQKFDNLYCC------RESI 297
Db        541 VAVMATGYGKSLCFQYPPVYVGKIGLVISPLISLMEDQVLQLKMSNIPACFLGSAQSENV 600

Qy        298 LDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEI---DPICALQ---AC 351
Db        601 L------TDIKL-GKYRIVYVTPEYCSGNMGLLQQLEADIGITLIAVDEAHCISEWGHDF 653

Qy        352 MDGFRVV-KLNEVIRQVDVV-ITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSL 409
Db        654 RDSFRKLGSLKTALPMVPIVALTATASSSI--REDIVRCLNL--------RNPQITCTGF 703

Qy        410 RTPELTWERVRSQVDHVI-------------WP-DGKRVV----------LLAEGRLLNL 445
Db        704 DRPNLYLE-VRRKTGNILQDLQPFLVKTSSHWEFEGPTIIYCPSRKMTQQVTGELRKLNL 762

Qy        446 SCSTVPT-------------FV---LSITATTQALALIELYNAPEGRYKQDV-----YLL 484
Db        763 SCGTYHAGMSFSTRKDIHHRFVRDEIQCVIATIAFGM--------GINKADIRQVIHYGA 814

Qy        485 PKKMDEY 491
Db        815 PKDMESY 821


RESULT 13
US-10-995-561-531
; Sequence 531, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 531
;   LENGTH: 1436
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-995-561-531

  Query Match             3.5%;  Score 96.5;  DB 6;  Length 1436;
  Best Local Similarity   21.1%;  Pred. No. 2.9;
  Matches  117;  Conservative   73;  Mismatches  199;  Indels  165;  Gaps   35;

Qy         36 KAPKKQIQFADDMQEFTKFPTKTGRRSLS---------RSISQSSTDSYSSAASYTDSSD 86
Db        366 KVERKEDGFEDGVED-NKLKENMERACLMSLDITEHELQILEQQSQEEYLSDIAY--KST 422

Qy         87 DEVSPREKQQTNS---KGSSNFCVKNIKQAEFGRREIEIAEQD--MSALISLR--KRAQG 139
Db        423 EHLSPNDNENDTSYVIESDEDLEMEMLKVCPVNNSPFVFGEKDTTLSVVLLLRFLLNFQS 482

Qy        140 EKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNI-YSTQNEVAAALAEAGVAVF 198
Db        483 LENL-NSGTVEPTH---------------SKCLKMERNLGLPTKEEEEDDENEAN---- 521

Qy        199 AWKGESED--DFWWCIDR-----CVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGI 251
Db        522 --EGEEDDDKDFLWPAPNEEQVTCLKM------------YFGHSSFK--PVQWKVIHSV 564

Qy        252 VEE-------SVTGVHR--LYQ-----LSKAGKLCVPAMN-VNDSVTKQKFDNLYCC--- 293
Db        565 LEERRDNVAVMATGYGKSLCFQYPPVYVGKIGLVISPLISLMEDQVLQLKMSNIPACFLG 624

Qy        294 ---RESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEI---DPICA 347
Db        625 SAQSENVL------TDIKL-GKYRIVYVTPEYCSGNMGLLQQLEADIGITLIAVDEAHCI 677

Qy        348 LQ---ACMDGFRVV-KLNEVIRQVDVV-ITCTGNKNVVTREHLDRMKNSCIVCNMGHSNT 402
Db        678 SEWGHDFRDSFRKLGSLKTALPMVPIVALTATASSSI--REDIVRCLNL--------RNP 727

Qy        403 EIDVTSLRTPELTWERVRSQVDHVI-------------WP-DGKRVV----------LLA 438
Db        728 QITCTGFDRPNLYLE-VRRKTGNILQDLQPFLVKTSSHWEFEGPTIIYCPSRKMTQQVTG 786

Qy        439 EGRLLNLSCSTVPT-------------FV---LSITATTQALALIELYNAPEGRYKQDV- 481
Db        787 ELRKLNLSCGTYHAGMSFSTRKDIHHRFVRDEIQCVIATIAFGM--------GINKADIR 838

Qy        482 ----YLLPKKMDEY 491
Db        839 QVIHYGAPKDMESY 852


RESULT 14
US-11-089-551A-28
; Sequence 28, Application US/11089551A
; Publication No. US20050266242A1
; GENERAL INFORMATION:
;  APPLICANT: Lindquist et al.
;  TITLE OF INVENTION: ELECTRICAL CONDUCTORS AND DEVICES FROM PRION-LIKE PROTEINS
;  FILE REFERENCE: 30554/40025A
;  CURRENT APPLICATION NUMBER: US/11/089,551A
;  CURRENT FILING DATE:  2005-03-24
;  PRIOR APPLICATION NUMBER: US 60/559,286
;  PRIOR FILING DATE: 2004-03-31
;  NUMBER OF SEQ ID NOS: 65
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 414
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-11-089-551A-28

  Query Match             3.4%;  Score 95;  DB 7;  Length 414;
  Best Local Similarity   25.5%;  Pred. No. 0.64;
  Matches   40;  Conservative   26;  Mismatches   67;  Indels   24;  Gaps    3;

Qy         17 KQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYS 76
Db         18 KQAKE-EKAKAVSSSSSESSSSSSSSSESESESESESSSSSSSSSDSESSSSSSSDSES 76

Qy         77 ----------SAASYTDSSDDEVSPREKQQT-----------NSKGSSNFCVKNIKQA 113
Db         77 EAETKKEESKDSSSSSSDSSSDEEEEEEKEETKKEESKESSSSDSSSSSSSDSESEKEES 136
```

```
Qy        114 EFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVG 150
              :|: | ||::     | :|:   |        ||
Db        137 NDKKRKSEDAEEEEDEESSNKKQKNEETEEPATIFVG 173


RESULT 15
US-11-034-569-6
; Sequence 6, Application US/11034569
; Publication No. US20050251876A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Aspartate Kinase
;  FILE REFERENCE: BB1430 PCT
;  CURRENT APPLICATION NUMBER: US/11/034,569
;  CURRENT FILING DATE:  2005-01-13
;  PRIOR APPLICATION NUMBER: US/09/890,813
;  PRIOR FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: PCT/US00/34396
;  PRIOR FILING DATE: 2000-12-19
;  PRIOR APPLICATION NUMBER: 60/172944
;  PRIOR FILING DATE: 1999-12-21
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
;   LENGTH: 560
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (168)
;   OTHER INFORMATION: Xaa=any amino acid
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (384)
;   OTHER INFORMATION: Xaa=any amino acid
US-11-034-569-6

  Query Match             3.2%;  Score 90;  DB 7;  Length 560;
  Best Local Similarity   19.0%;  Pred. No. 2.7;
  Matches  105;  Conservative   80;  Mismatches  224;  Indels  144;  Gaps   23;

Qy         54 FPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD----------------DEVSPREKQQ 96
              | |:|| |  :| :|    || |  |   | |                  |::|   |
Db         32 FGTRTGPRG-ARGLSMVVADSTSRRAKQADGGDGVLGAPVLGGLGMEGLGDQLSVVMKFG 90

Qy         97 TNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGC--THI 154
               :|  |:       : ||     :   |:    ::|   :      ||| | |||   |:
Db         91 GSSVSSA------ARMAEVAGLILTFPEERPVVVLSAMGKTTNNLLLAGEKAVGCGVIHV 144

Qy        155 T------AQTAVLIETLCALGAQCRWSACN--IYSTQNEVAAALAEAGVAVFAWKGESED 206
              :         :: |:|:  ||        ||  :|  :  :       |:|:
Db        145 SEIEEWNMVKSLHIKTVDELGLP---XICNTSLYELEQLL------KGIAMMKELTPRTS 195

Qy        207 DFWWCIDRCVN----------------MDGWQANMILDDGGDLTHWVYKKYPNVFKKIR 249
              |:    |::                | :   |  |       :   || | |::
Db        196 DYLVSFGECMSTRIFSAYLNKIRVKARQYDAFDIGFITTDEFGNADILEATYPAVAKRLH 255

Qy        250 G--IVEES---VTGVHRLYQLSKAGKLCV-------------PAMNVNDSVTKQKFDNL 290
              |  | : :   |||   | :  |:| :                |: : :    :  | :
Db        256 GDWIQDPAIPVVTGF--LGKGWKSGAVTTLGRGGSDLTATTIGKALGLREIQVWKDVDGV 313

Qy        291 YCCRESILDGLK--------RTTDVMFGGKQV----------------VVCGYGEVGKG 325
                |  :|    |           |:: : | ||                |   |
Db        314 LTCDPNIYPHAKTVPYLTFEEATELAYFGAQVLHPQSMRPAREGDIPVRVKNSYNP---- 369

Qy        326 CCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHL 385
                  || | ::           |  ||   ||  : |::    ::      : :     |
Db        370 -----KAPGTLI---------TRQRDMDXGLVVLTSIVLKSNVTMLDIVSTRMLGQYGFL 415

Qy        386 DRMKNSCIVCNMGHS-----NTEIDVTSLRTPELTWER----VRSQVDHVIWPDGK-RVV 435
               |:   | : ::  |      :|: |:    |  | |      |::|||:    |  :|
Db        416 ARVSGICYIEDLCISVDCVATSEVSVSVSLDPSKIWSRELIQQASELDHVVEELEKIAIV 475

Qy        436 LLAEGRLL-----NLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMD- 489
              | : | :      |:  |::   :|  |       : : :    :|  | :: |:    |
Db        476 RLLQQRAIISLIGNVEQSSL---ILEKTGRVLRKSGVNVQMISQGASKVNMSLIVHDSDA 532

Qy        490 -EYVASLHLPSFD 501
               | :||   |:
Db        533 KALVEALHQAFFE 545


Search completed: December 14, 2005, 08:31:42
Job time : 9.35962 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:11:12 ; Search time 32.6025 Seconds
(without alignments)
1564.139 Million cell updates/sec

Title: US-10-683-610-1
Perfect score: 2781
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....QAKYLGLNKNGPFKPNYYRY 530

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2582	92.8	597	2	T08681	adenosylhomocystei
2	1781.5	64.1	527	1	S01302	adenosylhomocystei
3	1184.5	42.6	433	1	JC2480	adenosylhomocystei
4	1156	41.6	432	1	A26583	adenosylhomocystei
5	1146	41.2	432	1	A43629	adenosylhomocystei
6	1142	41.1	466	2	AF2580	S-adenosylhomocyst
7	1142	41.1	466	2	D97362	adenosylhomocystei
8	1141	41.0	437	2	T32918	adenosylhomocystei
9	1130	40.6	437	1	A45569	adenosylhomocystei
10	1129	40.6	433	2	T40763	adenosylhomocystei
11	1121.5	40.3	481	2	AG3505	adenosylhomocystei
12	1110	39.9	485	2	S38379	adenosylhomocystei
13	1108.5	39.9	430	1	A27655	adenosylhomocystei
14	1108	39.8	485	2	C71400	adenosylhomocystei
15	1106	39.8	463	2	H87280	adenosylhomocystei
16	1105	39.7	449	2	S50546	adenosylhomocystei
17	1104	39.7	485	2	S71621	adenosylhomocystei
18	1103	39.7	485	2	T06764	adenosylhomocystei
19	1078.5	38.8	469	2	H83591	S-adenosyl-L-homoc
20	1072.5	38.6	462	1	A46035	adenosylhomocystei
21	1072	38.5	479	1	A54040	adenosylhomocystei
22	1052.5	37.8	495	2	B70593	adenosylhomocystei
23	1040.5	37.4	492	2	D87005	probable S-adenosy
24	1026	36.9	446	2	D82730	adenosylhomocystei
25	790.5	28.4	399	2	B72649	probable adenosylh
26	772.5	27.8	417	2	S58193	adenosylhomocystei
27	772.5	27.8	442	2	C90224	s-adenosyl-L-homoc
28	739	26.6	418	2	E70401	S-adenosylhomocyst
29	709	25.5	425	2	S75518	adenosylhomocystei
30	692	24.9	415	1	C64473	adenosylhomocystei
31	690	24.8	425	2	AC1983	adenosylhomocystei
32	684.5	24.6	417	2	F69085	adenosylhomocystei
33	674.5	24.3	422	2	B75064	adenosylhomocystei
34	674	24.2	405	2	G69499	adenosylhomocystei
35	672.5	24.2	425	2	H71167	probable S-adenosy
36	611.5	22.0	404	2	B72409	adenosylhomocystei
37	598.5	21.5	427	2	G84375	adenosylhomocystei
38	553	19.9	227	2	T15035	adenosylhomocystei
39	275	9.9	120	2	S22958	adenosylhomocystei
40	237.5	8.5	326	2	F69360	adenosylhomocystei
41	135.5	4.9	324	2	D97262	probable phosphogl
42	127.5	4.6	641	2	C82206	methyl-accepting c
43	121.5	4.4	515	2	I40809	glutamyl-tRNA redu
44	116	4.2	409	2	C82072	D-3-phosphoglycera
45	113	4.1	1506	2	T30886	integumentary muci

ALIGNMENTS

RESULT 1
T08681
adenosylhomocysteinase (EC 3.3.1.1) DKFZp564A1523 - human (fragment)
N;Alternate names: protein DKFZp564A1523.1
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08681
R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16469
A;Accession: T08681
A;Molecule type: mRNA
A;Residues: 1-597 <WAM>
A;Cross-references: UNIPROT:Q9UG84; UNIPARC:UPI000006FAA4; EMBL:AL049954
A;Experimental source: fetal brain; clone DKFZp564A1523
C;Genetics:
A;Note: DKFZp564A1523.1
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: NAD; thioether hydrolase
F;379-408/Region: beta-alpha-beta NAD nucleotide-binding fold
F;244,278/Active site: Cys #status predicted

Query Match 92.8%; Score 2582; DB 2; Length 597;
Best Local Similarity 88.8%; Pred. No. 2.7e-192;
Matches 502; Conservative 8; Mismatches 19; Indels 36; Gaps 4;

```
Qy          1 MSMPDAMPLP-------GVGE---ELKQAKEIEDAEKYSFMATVTKAPKK---------- 40
              :| | |:|||       | |    |:| :  | ::    :  :  |
Db         34 ISQP-ALPLPQETQAPSGESENLRNLRQGRRAESSQNLLSLRRLVSAQGNGRSDWRWLWI 92

Qy         41 --------------QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSS 85
                           ||||||||||||||||||||||||||||||||||||||||||||||
Db         93 QLWPEIDSLEISCFLQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSS 152

Qy         86 DDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAG 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 DDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAG 212

Qy        146 AKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESE 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 AKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESE 272

Qy        206 DDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQL 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 DDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQL 332

Qy        266 SKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKG 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 SKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKG 392

Qy        326 CCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHL 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        393 CCAALKALGAIVYITEIDPICALQACMDGFRVVVKLNEVIRQVDVVITCTGNKNVVTREHL 452

Qy        386 DRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNL 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 DRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNL 512

Qy        446 SCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLT 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 SCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLT 572

Qy        506 ELTDDQAKYLGLNKNGPFKPNYYRY 530
              |||||||||||||||||||||||||
Db        573 ELTDDQAKYLGLNKNGPFKPNYYRY 597


RESULT 2
S01302
adenosylhomocysteinase (EC 3.3.1.1) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S01302
R;DeLorenzi, M.; Ali, N.; Saari, G.; Henry, C.; Wilcox, M.; Bienz, M.
EMBO J. 7, 3223-3231, 1988
A;Title: Evidence that the Abdominal-B r element function is conferred by a trans-regula
A;Reference number: S01302; MUID:89030618; PMID:2903049
A;Accession: S01302
A;Molecule type: DNA
A;Residues: 1-527 <DEL>
A;Cross-references: UNIPROT:P50245; UNIPARC:UPI000016BB0E; EMBL:X13168; NID:g7675; PIDN:
C;Genetics:
A;Gene: FlyBase:Abd-B
A;Cross-references: FlyBase:FBgn0000015
A;Introns: 17/3; 41/1; 263/2; 369/3; 451/1
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: NAD; thioether hydrolase
F;275-304/Region: beta-alpha-beta NAD nucleotide-binding fold
F;140,174/Active site: Cys #status predicted

  Query Match             64.1%;  Score 1781.5;  DB 1;  Length 527;
  Best Local Similarity   70.8%;  Pred. No. 3.1e-130;
  Matches  328;  Conservative   64;  Mismatches   68;  Indels    3;  Gaps    3;

Qy         64 SRSISQSSTDSYSSAASYTDSSDDE-VSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEI 122
              :|  |  || |:||   :| |||:| |||::  | || | ::||||:| :: ||||||||
Db         27 ARRFSDVSTCSFSSTC-FTGSSDEEDVSPKDNHQRNSAGGTDFCVKSISKSAFGRREIEI 85

Qy        123 AEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYST 182
              || :|  :::||||: |||| || |||||: ||:||||| |  |||  ||:|||||||
Db         86 AESEMPGIMTLRKRAKDEKPLKGANIVGCTHVNAQSAVLIEALVQLGATVRWAACNIYST 145

Qy        183 QNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYP 242
              || ||||||||: :|||:||:|::||||:|| :  |||| |:|||||| || : ||||
Db        146 QNAVAAALAEAGIPIFAWRGETEEEFWWCLDRAIYSDGWQPNLILDDGGDATHLMLKKYP 205

Qy        243 NVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLK 302
              : || |||||||||||||| ||| ||| |||:|||||||| |||  | ||:|||| ||
Db        206 DYFKAIRGIVEESVTGVHRLYMLSKGGKLTVPAINVNDSVTKNKFDTFYTCRDSILDSLK 265

Qy        303 RTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNE 362
              ||||:|||||||:||||:|||||  :||  | |||:||:|||||||| |||||||:|||
Db        266 RTTDIMFGGKQVVICGYGDVGKGCAQSLKGQGCIVYVTEVDPICALQAAMDGFRVVRLNE 325

Qy        363 VIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQ 422
              ||| ||||:| ||||||:||:|::|||| ||:|||||| :||||  | ||||||||||||
Db        326 VIRTVDVVVTATGNKNVITRDHMNRMKNGCILCNMGHSCSEIDVNGLHTPELTWERVRSQ 385

Qy        423 VDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVY 482
              |||: |||||: ::|||||||:|||||||: :||:|: ::|||||||||::|| |||| |||
Db        386 VDHIRWPDGRMIILLAEGRLVNLSCSTISSFVVSVASSTQALALIELFSAP-GRYKSDVY 444

Qy        483 LLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKP 525
              |||||||||||||| :|||||||||:|:|::|||| |  ||
Db        445 LLPKKMDEYVASLHLATFDAHLTELTDEQSKFMGLNKAGLLKP 487


RESULT 3
JC2480
adenosylhomocysteinase (EC 3.3.1.1) - African clawed frog
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Xenopus laevis (African clawed frog)
C;Date: 29-Mar-1995 #sequence_revision 04-Oct-1996 #text_change 09-Jul-2004
C;Accession: JC2480
R;Seery, L.T.; McCabe, B.D.; Schoenberg, D.R.; Whitehead, A.S.
Biochem. Biophys. Res. Commun. 205, 1539-1546, 1994
A;Title: S-Adenosyl-L-homocysteine hydrolase from Xenopus laevis:Identification, devel
A;Reference number: JC2480; MUID:95110290; PMID:7811234
A;Accession: JC2480
A;Molecule type: mRNA
A;Residues: 1-433 <SEE>
A;Cross-references: UNIPROT:P51893; UNIPARC:UPI00001354E1; GB:L35559; NID:g558507; PID
A;Experimental source: liver
C;Function:
A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to ade
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: NAD; thioether hydrolase
F;215-244/Region: beta-alpha-beta NAD nucleotide-binding fold
F;79,113/Active site: Cys #status predicted

  Query Match             42.6%;  Score 1184.5;  DB 1;  Length 433;
  Best Local Similarity   52.9%;  Pred. No. 5e-84;
  Matches  227;  Conservative   81;  Mismatches  118;  Indels    3;  Gaps    3;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163
              :: | :|  |:||: ||||| :|  |: :|:     ||| ||:| || |:| |||||||
Db          6 SYKVADISLADWGRKAIEIAENEMPGLMKMREMHSESKPLKGARIAGCLHMTLQTAVLIE 65

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNM-DGWQ 222
              || ||||: :||:|||:|||:  |||:|: || |:|||||:::: |||:: :   ||
Db         66 TLTALGAEVQWSSCNIFSTQDHAAAAIAKTGVPVYAWKGETDEEYIWCIEQTIYFKDGKP 125

Qy        223 ANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSV 282
              ||||||||||: |: ||| : | |:|| ||: |||| ||::  :| | |||:||||||
Db        126 LNMILDDGGDLTNLVHSKYPQLLKGIKGISEETTTGVHNLYKMKSSGTLQVPAINVNDSV 185

Qy        283 TKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEI 342
              || |||||| ||||::||:|| ||||  ||  || |||:|||||  ||:| || | ||||
Db        186 TKSKFDNLYGCRESLIDGIKRATDVMIAGKVAVVAGYGDVGKGCAQALRAFGARVIITEI 245

Qy        343 DPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNT 402
              ||| |||| |:|: |  ::|  :: :: :| ||  ::|   | : ||:  ||||:|| :
Db        246 DPINALQAAMEGYEVTTMDEASKEGNIFVTTTGCADIVEGRHFENMKDDSIVCNIGHFDI 305

Qy        403 EIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATT 461
              |:||  |    :    :: ||| :  :|: ::|||||||:|| |:   |:||:| : |
Db        306 ELDVKWLNENAVKKVNIKPQVDRYLLKNGRHIILLAEGRLVNLGCAMGHPSFVMSNSFTN 365

Qy        462 QALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNG 521
              | :| |||:   : :|   || ||||:|| ||: ||       ||:||| |||||||:| |
Db        366 QVMAQIELWTNTD-KYPVGVYFLPKKLDEAVAAAHLDKLGVKLTKLTDKQAKYLGLDKEG 424

Qy        522 PFKPNYYRY 530
              ||||::|||
Db        425 PFKPDHYRY 433


RESULT 4
A26583
adenosylhomocysteinase (EC 3.3.1.1) - rat
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-Nov-1988 #sequence_revision 04-Oct-1996 #text_change 09-Jul-2004
C;Accession: A26583; A26191; B26191; S69333
R;Ogawa, H.; Gomi, T.; Mueckler, M.M.; Fujioka, M.; Backlund Jr., P.S.; Aksamit, R.R.;
Proc. Natl. Acad. Sci. U.S.A. 84, 719-723, 1987
A;Title: Amino acid sequence of S-adenosyl-L-homocysteine hydrolase from rat liver as
A;Reference number: A26583; MUID:87118240; PMID:3027698
```

A;Accession: A26583
A;Molecule type: mRNA
A;Residues: 1-432 <OGA>
A;Cross-references: UNIPROT:P10760; UNIPARC:UPI0000167A53; GB:M15185; NID:g202803; PIDN:
R;Gomi, T.; Ogawa, H.; Fujioka, M.
J. Biol. Chem. 261, 13422-13425, 1986
A;Title: S-Adenosylhomocysteinase from rat liver. Amino acid sequences of the peptides c
A;Reference number: A92560; MUID:87008564; PMID:3759971
A;Accession: A26191
A;Molecule type: protein
A;Residues: 76-94 <GOM>
A;Cross-references: UNIPARC:UPI0000172A00
A;Accession: B26191
A;Molecule type: protein
A;Residues: 104-121 <GO2>
A;Cross-references: UNIPARC:UPI0000172A00
A;Note: in a manner saturably inhibited by adenosine (a product) or adenine (a competiti
-79 and Cys-113, suggesting a role for these residues at the active site; however, neith
R;Merta, A.; Aksamit, R.R.; Kasir, J.; Cantoni, G.L.
Eur. J. Biochem. 229, 575-582, 1995
A;Title: The gene and pseudogenes of rat S-adenosyl-L-homocysteine hydrolase.
A;Reference number: S69333; MUID:95262723; PMID:7744082
A;Accession: S69333
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-432 <MER>
A;Cross-references: UNIPARC:UPI0000167A53; EMBL:U14937; NID:g1223843; PIDN:AAA92043.1; P
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1994
C;Genetics:
A;Introns: 10/1; 73/3; 99/1; 149/1; 186/3; 256/1; 285/2; 324/3; 389/3
C;Function:
A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to adenc
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: homotetramer; NAD; thioether hydrolase
F;214-243/Region: beta-alpha-beta NAD nucleotide-binding fold
F;79,113/Active site: Cys #status predicted

Query Match 41.6%; Score 1156; DB 1; Length 432;
Best Local Similarity 51.5%; Pred. No. 8.2e-82;
Matches 219; Conservative 84; Mismatches 120; Indels 2; Gaps 2;

```
Qy        107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
              | :|  | :||: ::||| :|  |: :|:     ||| ||:| || |:| :||||||||
Db          9 VADIGLAAWGRKALDIAENEMPGLMRMREMYSASKPLKGARIAGCLHMTVETAVLIETLV 68

Qy        167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMI 226
              ||||: |||:|||:|||:  |||:|:||: |||||||:::::: |||:: ::       |||
Db         69 ALGAEVRWSSCNIFSTQDHAAAAIAKAGIPVFAWKGETDEEYLWCIEQTLHFKDGPLNMI 128

Qy        227 LDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQK 286
              ||||||||: :: |:| :   |||| ||: |||| ||::   | | |||:|||||||| |
Db        129 LDDGGDLTNLIHTKHPQLLSGIRGISEETTTGVHNLYKMMANGILKVPAINVNDSVTKSK 188

Qy        287 FDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPIC 346
              ||||| |||||::||:|| ||||  ||  || |||:|||||  ||:  || | |||||||
Db        189 FDNLYGCRESLIDGIKRATDVMIAGKVAVVAGYGDVGKGCAQALRGFGARVIITEIDPIN 248

Qy        347 ALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDV 406
              |||| |:|: |  ::|  :: :: :| ||  :::   | ::||:  ||||:|| : ||||
Db        249 ALQAAMEGYEVTTMDEACKEGNIFVTTTGCVDIILGRHFEQMKDDAIVCNIGHFDVEIDV 308

Qy        407 TSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALA 465
                |    :   :: ||| :  :| |::||||||:|| |:   |:||:| : | | :|
Db        309 KWLNENAVEKVNIKPQVDRYLLKNGHRIILLAEGRLVNLGCAMGHPSFVMSNSFTNQVMA 368

Qy        466 LIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKP 525
              |||:  |: :|   |: ||||:|| ||  ||   :  ||:||: ||:|||:  ||||||
Db        369 QIELWTHPD-KYPVGVHFLPKKLDEAVAEAHLGKLNVKLTKLTEKQAQYLGMPINGPFKP 427

Qy        526 NYYRY 530
              ::|||
Db        428 DHYRY 432
```

RESULT 5
A43629
adenosylhomocysteinase (EC 3.3.1.1) - human
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Homo sapiens (man)
C;Date: 11-Feb-1993 #sequence_revision 04-Oct-1996 #text_change 09-Jul-2004
C;Accession: A43629; A60446; A57486; S65894
R;Coulter-Karis, D.E.; Hershfield, M.S.
Ann. Hum. Genet. 53, 169-175, 1989
A;Title: Sequence of full length cDNA for human S-adenosylhomocysteine hydrolase.
A;Reference number: A43629; MUID:90087640; PMID:2596825
A;Accession: A43629
A;Molecule type: mRNA
A;Residues: 1-432 <COU>
A;Cross-references: UNIPROT:P23526; UNIPARC:UPI00001354F1; GB:M61831
R;Arredondo-Vega, F.X.; Charlton, J.A.; Edwards, Y.H.; Hopkinson, D.A.; Whitehouse, D.
Ann. Hum. Genet. 53, 157-167, 1989
A;Title: Isozyme and DNA analysis of human S-adenosyl-L-homocysteine hydrolase (AHCY).
A;Reference number: A60446; MUID:90087639; PMID:2574561
A;Accession: A60446
A;Molecule type: mRNA
A;Residues: 122-432 <ARR>
A;Cross-references: UNIPARC:UPI00001729FA
A;Note: the authors translated the codon CTT for residue 162 as Glu, TAC for residue 2
R;Yuan, C.S.; Borchardt, R.T.
J. Biol. Chem. 270, 16140-16146, 1995
A;Title: Photoaffinity labeling of human placental S-adenosylhomocysteine hydrolase wi
A;Reference number: A57486; MUID:95332317; PMID:7608178
A;Accession: A57486
A;Status: preliminary
A;Molecule type: protein
A;Residues: 175-178,'X',180,'X',182-184,'X',186;319-326,'X' <YUA>
A;Cross-references: UNIPARC:UPI00001729FB; UNIPARC:UPI00001729FC
R;Gupta, R.A.; Yuan, C.S.; Ault-Riche, D.B.; Borchardt, R.T.
Arch. Biochem. Biophys. 319, 365-371, 1995
A;Title: Limited proteolysis of S-adenosylhomocysteine hydrolase: implications for the
A;Reference number: S65894; MUID:95305573; PMID:7786017
A;Accession: S65894
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-6;104-108;198-202 <GUP>
A;Cross-references: UNIPARC:UPI00001729FD; UNIPARC:UPI00001729FE; UNIPARC:UPI00001729F
C;Genetics:
A;Gene: GDB:AHCY
A;Cross-references: GDB:118983; OMIM:180960
A;Map position: 20cen-20q13.1
C;Function:
A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to ade
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: homotetramer; NAD; thioether hydrolase
F;214-243/Region: beta-alpha-beta NAD nucleotide-binding fold
F;79,113/Active site: Cys #status predicted

Query Match 41.2%; Score 1146; DB 1; Length 432;
Best Local Similarity 51.1%; Pred. No. 4.9e-81;
Matches 217; Conservative 86; Mismatches 120; Indels 2; Gaps 2;

```
Qy        107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
              | :|  | :||: ::||| :|  |: :|:|     ||| ||:| || |:| :||||||||
Db          9 VADIGLAAWGRKALDIAENEMPGLMRMRERYSASKPLKGARIAGCLHMTVETAVLIETLV 68

Qy        167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMI 226
              |||: :||:|||:|||:  |||:|:||: |:|||||:::::: |||:: :       |||
Db         69 TLGAEVQWSSCNIFSTQDHAAAAIAKAGIPVYAWKGETDEEYLWCIEQTLYFKDGPLNMI 128

Qy        227 LDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQK 286
              ||||||||: :: |:| :   |||| ||: |||| ||::   | | |||:|||||||| |
Db        129 LDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVHNLYKMMANGILKVPAINVNDSVTKSK 188

Qy        287 FDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPIC 346
              ||||| |||||::||:|| ||||  ||  || |||:|||||  ||:  || | |||||||
Db        189 FDNLYGCRESLIDGIKRATDVMIAGKVAVVAGYGDVGKGCAQALRGFGARVIITEIDPIN 248
```

```
Qy        347 ALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDV 406
              |||| |:|: |  ::|  :: :: :| ||  :::   | ::||:  ||||:|| : ||||
Db        249 ALQAAMEGYEVTTMDEACQEGNIFVTTTGCIDIILGRHFEQMKDDAIVCNIGHFDVEIDV 308

Qy        407 TSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALA 465
                |    :    :: |||     :|:|::|||||||:|| |:   |:||:| : | | :|
Db        309 KWLNENAVEKVNIKPQVDRYRLKNGRRIILLAEGRLVNLGCAMGHPSFVMSNSFTNQVMA 368

Qy        466 LIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKP 525
              |||:  |: :|   |: ||||:|| ||  ||   :  ||:||: ||:|||:: :|||||
Db        369 QIELWTHPD-KYPVGVHFLPKKLDEAVAEAHLGKLNVKLTKLTEKQAQYLGMSCDGPFKP 427

Qy        526 NYYRY 530
              ::|||
Db        428 DHYRY 432


RESULT 6
AF2580
S-adenosylhomocysteine hydrolase ahcY [imported] - Agrobacterium tumefaciens (strain C58
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AF2580
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2580
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <KUR>
A;Cross-references: UNIPROT:Q8UJ99; UNIPARC:UPI00001354E6; GB:AE008688; PIDN:AAL41060.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: ahcY
A;Map position: circular chromosome
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type

  Query Match           41.1%;  Score 1142;  DB 2;  Length 466;
  Best Local Similarity   50.2%;  Pred. No. 1.1e-80;
  Matches  233;  Conservative   68;  Mismatches  123;  Indels   40;  Gaps    6;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163
              :: | :|  | |||:|::||| :|  |:| ||     ||| ||:| |  |:| || ||||
Db          6 DYIVADINLAAFGRKELDIAETEMPGLMSCRKEFGESKPLKGARITGSLHMTIQTGVLIE 65

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVN-MDGWQ 222
              ||  |||: ||::|||:|||:  |||:| ||: ||| |||| ::|   |:     ||
Db         66 TLKELGAEIRWASCNIFSTQDHAAAAIAAAGIPVFAVKGESLTEYWEYTDKIFQWTDGGL 125

Qy        223 ANMILDDGGDLTHWV------------------------------YKKYPNVFKK---- 247
              :|||||||| | ::                              |  |  | |
Db        126 SNMILDDGGDATMYILLGARAEAGEDVLSNPGSEEEEILFAQINKRLKASPGWFTKQRDA 185

Qy        248 IRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDV 307
              ::|: ||: ||||||| |:| | |  ||:|||||||| |||| | |:||::||::| |||
Db        186 LKGVTEETTTGVHRLYDLAKKGLLPFPAINVNDSVTKSKFDNKYGCKESLVDGIRRATDV 245

Qy        308 MFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQV 367
              |  ||  ||||||:|||| |:|:  || | :|||||||||| |||| ||:| :|:
Db        246 MMAGKVAVVCGYGDVGKGSAASLQGAGARVKVTEIDPICALQAAMDGFEVVRLEDVVSSA 305

Qy        368 DVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVI 427
              |: || ||||:|:  ||:  || |:|| : || | |||   | |  :: ||| :
Db        306 DIFITTTGNKDVIRIEHMREMKDMAIVGNIGHFDNEIQVASLR--NLKWTNIKPQVDMIE 363

Qy        428 WPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPK 486
              :| | |::||:|||||||  :|  |:||:| : | | |  |||:  | | || :||:|||
Db        364 FPKGNRIILLSEGRLLNLGNATGHPSFVMSASFTNQVLGQIELFTKP-GEYKNEVYVLPK 422

Qy        487 KMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              :|| || |||       ||||:| || |:|::| ||||  :|||
Db        423 HLDEKVARLHLEKLGVRLTELSDLQADYIGISKQGPFKAEHYRY 466


RESULT 7
D97362
adenosylhomocysteinase (S-adenosyl-l-homocysteine hydrolase) (adohcyase) [imported] - ;
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: D97362
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldma
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium t
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: D97362
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <KUR>
A;Cross-references: UNIPROT:Q8UJ99; UNIPARC:UPI00001354E6; GB:AE007869; PIDN:AAK85853.
C;Genetics:
A;Gene: AGR_C_46
A;Map position: circular chromosome
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type

  Query Match           41.1%;  Score 1142;  DB 2;  Length 466;
  Best Local Similarity   50.2%;  Pred. No. 1.1e-80;
  Matches  233;  Conservative   68;  Mismatches  123;  Indels   40;  Gaps    6;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163
              :: | :|  | |||:|::||| :|  |:| ||     ||| ||:| |  |:| || ||||
Db          6 DYIVADINLAAFGRKELDIAETEMPGLMSCRKEFGESKPLKGARITGSLHMTIQTGVLIE 65

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVN-MDGWQ 222
              ||  |||: ||::|||:|||:  |||:| ||: ||| |||| ::|   |:     ||
Db         66 TLKELGAEIRWASCNIFSTQDHAAAAIAAAGIPVFAVKGESLTEYWEYTDKIFQWTDGGL 125

Qy        223 ANMILDDGGDLTHWV------------------------------YKKYPNVFKK---- 247
              :|||||||| | ::                              |  |  | |
Db        126 SNMILDDGGDATMYILLGARAEAGEDVLSNPGSEEEEILFAQINKRLKASPGWFTKQRDA 185

Qy        248 IRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDV 307
              ::|: ||: ||||||| |:| | |  ||:|||||||| |||| | |:||::||::| |||
Db        186 LKGVTEETTTGVHRLYDLAKKGLLPFPAINVNDSVTKSKFDNKYGCKESLVDGIRRATDV 245

Qy        308 MFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQV 367
              |  ||  ||||||:|||| |:|:  || | :|||||||||| |||| ||:| :|:
Db        246 MMAGKVAVVCGYGDVGKGSAASLQGAGARVKVTEIDPICALQAAMDGFEVVRLEDVVSSA 305

Qy        368 DVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVI 427
              |: || ||||:|:  ||:  || |:|| : || | |||   | |  :: ||| :
Db        306 DIFITTTGNKDVIRIEHMREMKDMAIVGNIGHFDNEIQVASLR--NLKWTNIKPQVDMIE 363

Qy        428 WPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPK 486
              :| | |::||:|||||||  :|  |:||:| : | | |  |||:  | | || :||:|||
Db        364 FPKGNRIILLSEGRLLNLGNATGHPSFVMSASFTNQVLGQIELFTKP-GEYKNEVYVLPK 422

Qy        487 KMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              :|| || |||       ||||:| || |:|::| ||||  :|||
Db        423 HLDEKVARLHLEKLGVRLTELSDLQADYIGISKQGPFKAEHYRY 466


RESULT 8
T32918
adenosylhomocysteinase (EC 3.3.1.1) - Caenorhabditis elegans
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32918; T41749
R;Maggi, L.; Goela, D.
```

submitted to the EMBL Data Library, January 1998
A;Description: The sequence of C. elegans cosmid K02F2.
A;Reference number: Z21246
A;Accession: T32918
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-437 <MAG>
A;Cross-references: UNIPROT:P27604; UNIPARC:UPI00001354EA; EMBL:AF043699; PIDN:AAB97565.
A;Experimental source: strain Bristol N2; clone K02F2
R;Prasad, S.S.; Starr, T.V.; Rose, A.M.
Genome 36, 57-65, 1993
A;Title: Molecular characterization in the dpy-14 region identifies the adedenosylhomocy
A;Reference number: Z22017; MUID:93209535; PMID:8458573
A;Accession: T41749
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-437 <PRA>
A;Cross-references: UNIPARC:UPI00001354EA; EMBL:S57284; PIDN:AAB25906.1
C;Genetics:
A;Gene: K02F2.2
A;Map position: 1
A;Introns: 74/3; 326/3
C;Function:
A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to adenc
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: NAD; thioether hydrolase

Query Match 41.0%; Score 1141; DB 2; Length 437;
Best Local Similarity 52.4%; Pred. No. 1.2e-80;
Matches 225; Conservative 78; Mismatches 120; Indels 6; Gaps 4;

```
Qy        107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
              | :|| |:|||:|| :|| :|  |:::| :    :|| ||:| || |:| |||||||||
Db         10 VADIKLADFGRKEIILAENEMPGLMAMRSKYGPSQPLKGARIAGCLHMTIQTAVLIETLT 69

Qy        167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDR-CVNMDGWQANM 225
              ||||: :||:|||:|||:  |||:|: || |:|||||::::: |||::  |  ||   ||
Db         70 ALGAEVQWSSCNIFSTQDHAAAAIAQTGVPVYAWKGETDEEYEWCIEQTIVFKDGQPLNM 129

Qy        226 ILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQ 285
              |||||||||: |: |||     |||: ||: |||| | ::   | | |||:||||||||
Db        130 ILDDGGDLTNLVHAKYPQYLAGIRGLSEETTTGVHNLAKMLAKGDLKVPAINVNDSVTKS 189

Qy        286 KFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPI 345
              ||||||  |||: ||:|| ||||  ||  || |||:|||| |:||| |: | :||||||
Db        190 KFDNLYGIRESLPDGIKRATDVMLAGKVAVVAGYGDVGKGSAASLKAFGSRVIVTEIDPI 249

Qy        346 CALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEID 405
              ||||  |:|: |  | |   : ::::| || |::|| :| : : |  ||||:|| : |||
Db        250 NALQAAMEGYEVTTLEEAAPKANIIVTTTGCKDIVTGKHFELLPNDAIVCNVGHFDCEID 309

Qy        406 VTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQAL 464
              |  | |      : :: |||     :|: |:|||||||:|| |:|  |:||:| : | | |
Db        310 VKWLNTNATKKDTIKPQVDRYTLKNGRHVILLAEGRLVNLGCATGHPSFVMSNSFTNQVL 369

Qy        465 ALIEL---YNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNG 521
              | :||   :  |:  ||  :|:||| :|| || |||      ||:|:|:|| |||:   |
Db        370 AQVELWTKFGTPQ-EYKLGLYVLPKTLDEEVAYLHLAQLGVKLTKLSDEQASYLGVPVAG 428

Qy        522 PFKPNYYRY 530
              |:||::|||
Db        429 PYKPDHYRY 437
```

RESULT 9
A45569
adenosylhomocysteinase (EC 3.3.1.1) - Leishmania donovani
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Leishmania donovani
C;Date: 03-Feb-1994 #sequence_revision 04-Oct-1996 #text_change 15-Mar-2004
C;Accession: A45569
R;Henderson, D.M.; Hanson, S.; Allen, T.; Wilson, K.; Coulter-Karis, D.E.; Greenberg, M.
Mol. Biochem. Parasitol. 53, 169-183, 1992
A;Title: Cloning of the gene encoding Leishmania donovani S-adenosylhomocysteine hydro
A;Reference number: A45569; MUID:92365726; PMID:1501636
A;Accession: A45569
A;Molecule type: DNA
A;Residues: 1-437 <HEN>
A;Cross-references: UNIPARC:UPI0000172A02
A;Note: this sequence is inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:111168, NCBIP:111170)
C;Function:
A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to ade
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: NAD; thioether hydrolase
F;213-242/Region: beta-alpha-beta NAD nucleotide-binding fold
F;76,110/Active site: Cys #status predicted

Query Match 40.6%; Score 1130; DB 1; Length 437;
Best Local Similarity 52.2%; Pred. No. 8.6e-80;
Matches 228; Conservative 74; Mismatches 125; Indels 10; Gaps 5;

```
Qy        103 SNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLI 162
              ::: ||:|  ||:||: ||:|| :|  |: ||:     :|| |||| || |:| ||||||
Db          2 ADYKVKDISLAEWGRKAIELAENEMPGLMELRREYGPSQPLKGAKIAGCLHMTVQTAVLI 61

Qy        163 ETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCV---NMD 219
              ||| ||||  |||:|||:|||:  |||:|:||| |||||||::::: ||| : |   : |
Db         62 ETLKALGADVRWSSCNIFSTQDNAAAAIAKAGVPVFAWKGETDEEYEWCIAQTVKGFSGD 121

Qy        220 GWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVN 279
              |   |||||||||||: |   :| :  || || ||: |||  ||:    | | : |:|||
Db        122 G-LPNMILDDGGDLTNLVIDHHPELVPKIFGISEETTTGVKNLYKRLSKGNLPMCAINVN 180

Qy        280 DSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYI 339
              ||||| |||||| ||||::||:|| ||||  ||   |||||:||||| |||:| || | :
Db        181 DSVTKSKFDNLYGCRESLVDGMKRATDVMIAGKTCCVCGYGDVGKGCAAALRAFGARVVV 240

Qy        340 TEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGH 399
              ||:||| |||| |:|:::|  : :|:    : :| ||| :::| ||   |::  ||||:||
Db        241 TEVDPINALQASMEGYQVALVEDVMADAHIFVTTTGNDDIITSEHFPHMRDDAIVCNIGH 300

Qy        400 SNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSC-STVPTFVLSIT 458
              :||| |  |          :: |||     :|: ::|||:|||:|| | |  |:||:| :
Db        301 FDTEIQVGWLEANAKEHVEIKPQVDFYTMENGRHIILLAKGRLVNLGCASGHPSFVMSNS 360

Qy        459 ATTQALALIELY-NAPEGRY----KQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAK 513
              | | || |||:  |   |:|    |  |: ||| :|| ||:||| |  |  |:||:| :
Db        361 FTNQVLAQIELWSNRDNGKYPRGDKAGVFFLPKALDEKVAALHLAHVGAKLTKLTPKQAE 420

Qy        514 YLGLNKNGPFKPNYYRY 530
              |:    ||||||::|||
Db        421 YINCPVNGPFKPDHYRY 437
```

RESULT 10
T40763
adenosylhomocysteinase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40763
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998
A;Reference number: Z21948
A;Accession: T40763
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-433 <LYN>
A;Cross-references: UNIPROT:O13639; UNIPARC:UPI000013550C; EMBL:AL022072; PIDN:CAA178
A;Experimental source: strain 972h-; cosmid c8D2
C;Genetics:
A;Gene: SPDB:SPBC8D2.18c
A;Map position: 2
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type

Query Match 40.6%; Score 1129; DB 2; Length 433;

```
  Best Local Similarity   50.7%;  Pred. No. 1e-79;
  Matches  219;  Conservative   81;  Mismatches  128;  Indels    4;  Gaps    3;

Qy        102 SSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVL 161
              : |: | :|  | |||:|:|||| :|  ||::|::    :|| ||:| || |:| |||||
Db          3 TQNYKVADISLAAFGRKELEIAENEMPGLIAVREKYAKSQPLKGARIAGCLHMTIQTAVL  62

Qy        162 IETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNM--D 219
              |||| ||||:  ||:|||||||:  |||:|  || |||||||:|::: |||:: :
Db         63 IETLVALGAEVTWSSCNIYSTQDHAAAAIAATGVPVFAWKGETEEEYLWCIEQQLKSFPS 122

Qy        220 GWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVN 279
              |   |||||||||||  |::::| :   |||| ||: |||| ||:: |  || |||:|||
Db        123 GKPLNMILDDGGDLTALVHERHPELLVDIRGISEETTTGVHNLYKMFKENKLKVPAINVN 182

Qy        280 DSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYI 339
              ||||| |||||: |:||::||:|| |||| || || |:|:|||||  :|:: || | :
Db        183 DSVTKSKFDNLFGCKESLVDGIKRATDVMIAGKVAVVAGFGDVGKGCSTSLRSQGARVIV 242

Qy        340 TEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGH 399
              ||:||| |||| |||| |  : | ::: : :| || :::: || : ||  ||||:||
Db        243 TEVDPINALQAAMDGFEVTTMEEAVKEGQIFVTTTGCRDIIRGEHFNEMKEDSIVCNIGH 302

Qy        400 SNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSIT 458
              : |||| |:        :: |||     :|: ::|||:|||:|| |:|  |:||:| :
Db        303 FDVEIDVAWLKANAKDVVNIKPQVDRYELKNGRHIILLADGRLVNLGCATGHPSFVMSCS 362

Qy        459 ATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLN 518
              | | || | |:     |   |::||||:|| || |||      || ||  |: |||:
Db        363 FTNQVLAQIALW-TDNTSYPLGVHMLPKKLDEEVARLHLGKLGVKLTTLTSVQSDYLGIP 421

Qy        519 KNGPFKPNYYRY 530
              :||:| ::|||
Db        422 VDGPYKADHYRY 433


RESULT 11
AG3505
adenosylhomocysteinase (EC 3.3.1.1) [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Mar-2004
C;Accession: AG3505
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AG3505
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-481 <KUR>
A;Cross-references: UNIPARC:UPI00000582D0; GB:AE008917; PIDN:AAL53210.1; PID:g17984086;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI2029
A;Map position: I
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: thioether hydrolase

  Query Match             40.3%;  Score 1121.5;  DB 2;  Length 481;
  Best Local Similarity   48.6%;  Pred. No. 4.5e-79;
  Matches  233;  Conservative   75;  Mismatches  124;  Indels   47;  Gaps    8;

Qy         96 QTNSKG-------SSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKI 148
              | |  |      | :| ||:|  |::||||::||| :|  |:: |:     :|| ||:|
Db          6 QANCPGLEHAMTASQDFVVKDISLADWGRRELDIAETEMPGLMAAREEFGKSQPLKGARI  65

Qy        149 VGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDF 208
              |  |:| |||||||||  |||: ||::|||:|||:  |  ||| |||: :::
Db         66 SGSLHMTIQTAVLIETLKVLGAEVRWASCNIFSTQDHAAAAIAATGTPVFAVKGETLEEY 125

Qy        209 WWCIDRCVNM-DGWQANMILDDGGDLTHWV----------------------------YKK 240
              |   |:     ||  :|||||||| | ::                              ||
Db        126 WTYTDQIFQWPDGEPSNMILDDGGDATMYILIGARAEAGEDVLSNPQSEEEEVLFAQIKK 185

Qy        241 ----YPNVFKK----IRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYC 292
                  | | |    |:|: ||: |||:|||| | | |  ||:|||||||| |||| |
Db        186 RMAATPGFFTKQRAAIKGVTEETTTGVNRLYQLQKKGLLPFPAINVNDSVTKSKFDNKYG 245

Qy        293 CRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACM 352
              |:||::||::| |||| ||  |||||:|||   :|  || | :||:|||||||| |
Db        246 CKESLVDGIRRGTDVMMAGKVAVVCGYGDVGKGSAQSLAGAGARVKVTEVDPICALQAAM 305

Qy        353 DGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTP 412
              ||| || |::     |:|:| ||||:|:| :|: :||: ||| |:|| : || | :||
Db        306 DGFEVVTLDDAASTADIVVTTTGNKDVITIDHMRKMKDMCIVGNIGHFDNEIQVAALR-- 363

Qy        413 ELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYN 471
              | |  |: ||| : :|||||::||:|||||||  :|  |:||:| : | | |  |||:
Db        364 NLKWTNVKPQVDLIEFPDGKRLILLSEGRLLNLGNATGHPSFVMSASFTNQVLGQIELFT 423

Qy        472 APEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              :  || :||:||| :|| || |||    | || |:::|| |:|:   ||||  :|||
Db        424 RTDA-YKNEVYVLPKHLDEKVARLHLDKLGAKLTVLSEEQAAYIGVTPQGPFKSEHYRY 481


RESULT 12
S38379
adenosylhomocysteinase (EC 3.3.1.1) - Madagascar periwinkle
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Catharanthus roseus (Madagascar periwinkle)
C;Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S38379
R;Schroeder, G.; Waitz, A.; Hotze, M.; Schroeder, J.
submitted to the EMBL Data Library, October 1993
A;Description: cDNA for stress-induced S-adenosyl-L-homocysteine hydrolase from Madaga
A;Reference number: S38379
A;Accession: S38379
A;Molecule type: DNA
A;Residues: 1-485 <SCH>
A;Cross-references: UNIPROT:P35007; UNIPARC:UPI00001354EB; EMBL:Z26881; NID:g407411; P
C;Genetics:
A;Gene: SAHH
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: NAD; thioether hydrolase
F;263-292/Region: beta-alpha-beta NAD nucleotide-binding fold
F;86,120/Active site: Cys #status predicted

  Query Match             39.9%;  Score 1110;  DB 2;  Length 485;
  Best Local Similarity   46.4%;  Pred. No. 3.6e-78;
  Matches  224;  Conservative   85;  Mismatches  124;  Indels   50;  Gaps    8;

Qy         95 QQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHI 154
              ::|:|     : ||:: ||:||| |||:|| :|  |:| |      :|  |||| |  |:
Db          6 EKTSS--GREYKVKDMSQADFGRLEIELAEVEMPGLMSCRAEFGPSQPFKGAKITGSLHM  63

Qy        155 TAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDR 214
              | ||||||||| |||: || :|||:|||   |||:|    ||||||||:  ::||| :|
Db         64 TIQTAVLIETLTALGAEVRWCSCNIFSTQEHAAAAIARDSAAVFAWKGETLQEYWWCTER 123

Qy        215 CVNM--DGWQANMILDDGGDLTHWVYK------------------------------- 239
              ::   ||   ::|:||||| | :::
Db        124 ALDWGPDG-GPDLIVDDGGDATLLIHEGVKAEEEYKKNGALPDPSSTDNAEFQIVLTIIR 182

Qy        240 --------KYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLY 291
                     ||  : ::: |: ||: ||| ||||:   | |  ||:|||||||| ||||||
Db        183 DGLKSDPTKYTRMKERLVGVSEETTTGVKRLYQMQANGTLLFPAINVNDSVTKSKFDNLY 242

Qy        292 CCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQAC 351
              || |: ||| | |||| ||  || |||:||||| ||||  || | :|||||||||||
Db        243 GCRHSLPDGLMRATDVMIAGKVAVVAGYGDVGKGCAAALKQAGARVIVTEIDPICALQAT 302

Qy        352 MDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRT 411
              |:| :|: | :|: : |: :| ||||:::  :|: :|||: ||||:|| : |||:  | |
Db        303 MEGLQVLTLEDVVSEADIFVTTTGNKDIIMVDHMRKMKNNAIVCNIGHFDNEIDMLGLET 362
```

```
Qy        412 -PELTWERVRSQVDHVIWPD-GKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIE 468
Db        363 YPGVKRITIKPQTDRWVFPDTNSGIIVLAEGRLMNLGCATGHPSFVMSCSFTNQVIAQLE 422

Qy        469 LYNAPE-GRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNY 527
Db        423 LWNERKTGKYEKKVYVLPKHLDEKVAALHLGKLGAKLTKLTKDQADYISVPIEGPYKPAH 482

Qy        528 YRY 530
Db        483 YRY 485


RESULT 13
A27655
adenosylhomocysteinase (EC 3.3.1.1) - slime mold (Dictyostelium discoideum)
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Dictyostelium discoideum
C;Date: 31-Mar-1989 #sequence_revision 04-Oct-1996 #text_change 09-Jul-2004
C;Accession: A27655; S06394
R;Kasir, J.; Aksamit, R.R.; Backlund Jr., P.S.; Cantoni, G.L.
Biochem. Biophys. Res. Commun. 153, 359-364, 1988
A;Title: Amino acid sequence of S-adenosyl-L-homocysteine hydrolase from Dictyostelium d
A;Reference number: A27655; MUID:88240426; PMID:3288206
A;Accession: A27655
A;Molecule type: mRNA
A;Residues: 1-430 <KAS>
A;Cross-references: UNIPROT:P10819; UNIPARC:UPI00001354EE; GB:M19937; NID:g167663; PIDN:
R;Guitton, M.C.; Part, D.; Veron, M.
Biochimie 70, 835-840, 1988
A;Title: Cloning of a cDNA for the S-adenosyl-L-homocysteine hydrolase from Dictyosteliu
A;Reference number: S06394; MUID:89000991; PMID:3139100
A;Accession: S06394
A;Molecule type: mRNA
A;Residues: 'VVTKSSPWKPPLHYQT',270-311,'G',313-323,'S',325-359,'SVTK',365-382,'F',384-43
A;Cross-references: UNIPARC:UPI000016B98D; EMBL:X12523; NID:g7204; PIDN:CAA31040.1; PID:
A;Note: this sequence differs at the amino end by an apparant frame shift; comparison to
C;Function:
A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to adeno
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: NAD; thioether hydrolase
F;213-242/Region: beta-alpha-beta NAD nucleotide-binding fold
F;78,112/Active site: Cys #status predicted

  Query Match             39.9%;  Score 1108.5;  DB 1;  Length 430;
  Best Local Similarity   51.9%;  Pred. No. 3.9e-78;
  Matches  222;  Conservative   72;  Mismatches  131;  Indels    3;  Gaps    3;

Qy        104 NFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIE 163
Db          5 HYKVKDISLAAWGRKEIEIAANEMPGLMTLRKKYGPAQILKGARIAGCLHMTIQTAVLIE 64

Qy        164 TLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQA 223
Db         65 TLTALGAQVQWSSCNIFSTQDQAAAAIAATGVPVYAWKGETEEEYNWCVEQTIVFQDGPL 124

Qy        224 NMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVT 283
Db        125 NMILDDGGDLTTLVHEKYPQFLAGIKGISEETTHGVHNLYKMFKEGKLKVPAINVNDSVT 184

Qy        284 KQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEID 343
Db        185 KSKFDNLYGCRESLIDGIKRATDVMIAGKVAVVAGYGDVGKGCAQSLSKMGARVLVTEID 244

Qy        344 PICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTE 403
Db        245 PINALQACMDGYQIVTMETAAPLSNIFVTTTGCRDIVRGEHFAVMKEDAIVCNIGHFDCE 304

Qy        404 IDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQ 462
Db        305 IDVAWLNA-NAKKDTVKPQVDRYTLANGVHIILLAEGRLVNLGCGTGHPSFVMSNSFCNQ 363

Qy        463 ALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGP 522
Db        364 TLAQIALWTKTE-EYPLGVHLLPKILDEEVARLHLDQLGAKLTTLTEKQSEYLSVPVAGP 422

Qy        523 FKPNYYRY 530
Db        423 YKVDHYRY 430


RESULT 14
C71400
adenosylhomocysteinase (EC 3.3.1.1) [similarity] - Arabidopsis thaliana
N;Alternate names: protein DL3010W; S-adenosyl-L-homocysteine hydrolase
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C;Accession: C71400; T51608
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Di
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Gi
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomene
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; A
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis th
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: C71400
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-485 <BEV>
A;Cross-references: UNIPROT:O23255; UNIPARC:UPI0000000A82; GB:Z97335; NID:g2244747; PI
R;Belbahri, L.; Elleuch, H.; Villaroel, R.; Inze, D.; Thomas, D.; Thomasset, B.
Plant Physiol. 121, 313, 1999
A;Title: The Isolation of an Arabidopsis thaliana cDNA clone encoding S-adenosyl-L-hom
A;Reference number: Z25409
A;Accession: T51608
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-485 <BEL>
A;Cross-references: UNIPARC:UPI0000000A82; EMBL:AF059581; PIDN:AAC14714.1
A;Experimental source: cultivar Landsberg erecta
C;Genetics:
A;Gene: SAHH
A;Map position: 4COP9-4G3845
A;Note: DL3010W
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: NAD; thioether hydrolase
F;263-292/Region: beta-alpha-beta NAD nucleotide-binding fold
F;86,120/Active site: Cys #status predicted

  Query Match             39.8%;  Score 1108;  DB 2;  Length 485;
  Best Local Similarity   45.2%;  Pred. No. 5.1e-78;
  Matches  218;  Conservative   92;  Mismatches  124;  Indels   48;  Gaps    7;

Qy         95 QQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHI 154
Db          6 EKTSS--GREYKVKDMSQADFGRLELELAEVEMPGLMACRTEFGPSQPFKGARITGSLHM 63

Qy        155 TAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDR 214
Db         64 TIQTAVLIETLTALGAEVRWCSCNIFSTQDHAAAAIARDSAAVFAWKGETLQEYWWCTER 123

Qy        215 CVNM-DGWQANMILDDGGDLTHWVY---------------------------------- 238
Db        124 ALDWGPGGGPDLIVDDGGDATLLIHEGVKAEEIFEKTGQVPDPTSTDNPEFQIVLSIIKE 183

Qy        239 ------KKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYC 292
Db        184 GLQVDPKKYHKMKERLVGVSEETTTGVKRLYQMQQNGTLLFPAINVNDSVTKSKFDNLYG 243

Qy        293 CRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACM 352
Db        244 CRHSLPDGLMRATDVMIAGKVAVICGYGDVGKGCAAAMKTAGARVIVTEIDPICALQALM 303
```

```
Qy        353 DGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRT- 411
              :| :|: | :|: : |: :| ||||::: :|: :|||: ||||:|| : |||:  | |
Db        304 EGLQVLTLEDVVSEADIFVTTTGNKDIIMVDHMRKMKNNAIVCNIGHFDNEIDMLGLETY 363

Qy        412 PELTWERVRSQVDHVIWPDGKR-VVLLAEGRLLNLSCST-VPTFVLSITATTQALALIEL 469
              | :    :: | |  ::|: |  :::|||||:|| |:|  |:||:| : | | :| :||
Db        364 PGVKRITIKPQTDRWVFPETKAGIIVLAEGRLMNLGCATGHPSFVMSCSFTNQVIAQLEL 423

Qy        470 YN-APEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYY 528
              :|    |:|:: ||:||| :|| || |||    | ||:|: ||: |: :   ||:|| :|
Db        424 WNEKASGKYEKKVYVLPKHLDEKVALLHLGKLGARLTKLSKDQSDYVSIPIEGPYKPPHY 483

Qy        529 RY 530
              ||
Db        484 RY 485


RESULT 15
H87280
adenosylhomocysteinase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: H87280
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: H87280
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-463 <STO>
A;Cross-references: UNIPROT:Q9ABH0; UNIPARC:UPI00001354EC; GB:AE005673; NID:g13421390; P
C;Genetics:
A;Gene: CC0257
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type

  Query Match             39.8%;  Score 1106;  DB 2;  Length 463;
  Best Local Similarity   47.9%;  Pred. No. 6.8e-78;
  Matches  226;  Conservative   82;  Mismatches  110;  Indels   54;  Gaps    8;

Qy        103 SNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLI 162
              ::: ||:|  |:|||:|| ||| :|  |:: |     :: | ||:| |  |:| ||||||
Db          2 ADYIVKDISLADFGRKEIAIAETEMPGLMATRAEYGPQQILKGARIAGSLHMTIQTAVLI 61

Qy        163 ETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFW--------WCIDR 214
              ||| ||||: ||::|||:|||:  |||:| ||| |||:|||:  ::|         |
Db         62 ETLTALGAEVRWASCNIFSTQDHAAAAIAAAGVPVFAFKGENLVEYWEYAHKIFEW---- 117

Qy        215 CVNMDGWQANMILDDGGDLT--------------------------HWVYKKY----PN 243
                 ||   |:|||||| |                            : | |||    |
Db        118 ---HDGGYPNLILDDGGDATLLCVLGPKAEKDPSILNNPQNEEEEALYAVMKKYLAEKPG 174

Qy        244 VFKKIR----GIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILD 299
              :  ||    |: ||: ||||||||:::  :|  ||:|||||||| |||||| ||||::|
Db        175 FYSAIRAAIGGVSEETTTGVHRLYQMAQKDELPFPAINVNDSVTKSKFDNLYGCRESLVD 234

Qy        300 GLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVK 359
              ::| ||||  ||  ||||||:|||| |:|:  || | :||:|||||||| |:|: |
Db        235 AIRRGTDVMLSGKVAVVCGYGDVGKGSAASLRQGGARVIVTEVDPICALQAAMEGYEVQT 294

Qy        360 LNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERV 419
              ||:|  : |: :| |||||:|:| : : :|||: ||||:|| ::|| :  ||      |: :
Db        295 LNDVADKADIFVTATGNKDVITVDDMRKMKNNAIVCNIGHFDSEIQIAGLR--NFKWDEI 352

Qy        420 RSQVDHVIWPDGKRVVLLAEGRLLNLSCST-VPTFVLSITATTQALALIELYNAPEGRYK 478
              : || || :||||:::::|:||||:||  :|  |:||:| : | | || |||:   : :|:
Db        353 KPQVHHVEFPDGKKLIVLSEGRLVNLGNATGHPSFVMSASFTNQTLAQIELW-TNKAKYE 411

Qy        479 QDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              || ||| :|| || |||    | || |  ||| |:|: : |||||::|||
Db        412 NQVYTLPKHLDEKVAFLHLEKLGAKLTTLRKDQADYIGVPEAGPFKPDHYRY 463
```

Search completed: December 14, 2005, 08:20:56
Job time : 34.6025 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:10:17 ; Search time 172.208 Seconds
(without alignments)
2171.383 Million cell updates/sec

Title: US-10-683-610-1
Perfect score: 2781
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....QAKYLGLNKNGPFKPNYYRY 530

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2781	100.0	530	2	Q96PK4_HUMAN	Q96pk4 homo sapien
2	2781	100.0	530	2	Q80SW1_MOUSE	Q80sw1 mus musculu
3	2781	100.0	607	2	Q502W8_HUMAN	Q502w8 homo sapien
4	2781	100.0	623	2	Q6P171_HUMAN	Q6p171 homo sapien
5	2627	94.5	500	1	SAHH2_HUMAN	O43865 homo sapien
6	2548	91.6	484	2	Q9UG84_HUMAN	Q9ug84 homo sapien
7	2542	91.4	483	2	Q5VSM0_HUMAN	Q5vsm0 homo sapien
8	2469	88.8	540	2	Q5F3Q3_CHICK	Q5f3q3 gallus gall
9	2466	88.7	512	2	Q7T2B5_BRARE	Q7t2b5 brachydanio
10	2459	88.4	520	2	Q66IP1_XENLA	Q66ip1 xenopus lae
11	2456	88.3	611	1	SAHH3_HUMAN	Q96hn2 homo sapien
12	2454	88.2	613	2	Q68FL4_MOUSE	Q68fl4 mus musculu
13	2446.5	88.0	588	2	Q5BL43_XENTR	Q5bl43 xenopus tro
14	2439	87.7	508	2	Q8BIH1_MOUSE	Q8bih1 mus musculu
15	2438.5	87.7	588	2	Q68EX1_XENLA	Q68ex1 xenopus lae
16	2433	87.5	491	2	Q4RXQ1_TETNG	Q4rxq1 tetraodon n
17	2430	87.4	508	2	Q5R889_PONPY	Q5r889 pongo pygma
18	2427	87.3	530	2	Q6ZNS6_HUMAN	Q6zns6 homo sapien
19	2410	86.7	591	2	Q6PC06_BRARE	Q6pc06 brachydanio
20	2408	86.6	583	2	Q6AZU3_XENLA	Q6azu3 xenopus lae
21	2313.5	83.2	585	2	Q4RVT5_TETNG	Q4rvt5 tetraodon n
22	2278	81.9	478	2	Q80TQ9_MOUSE	Q80tq9 mus musculu
23	2182	78.5	541	2	Q4RHJ7_TETNG	Q4rhj7 tetraodon n
24	1970	70.8	521	2	Q9VZX9_DROME	Q9vzx9 drosophila
25	1966	70.7	521	2	Q6NR07_DROME	Q6nr07 drosophila
26	1826.5	65.7	492	1	SAHH2_DROME	P50245 drosophila
27	1819.5	65.4	504	2	Q7KSG0_DROME	Q7ksg0 drosophila
28	1818	65.4	432	2	Q7QE71_ANOGA	Q7qe71 anopheles g
29	1763	63.4	522	2	Q5DH82_SCHJA	Q5dh82 schistosoma
30	1729	62.2	499	2	Q4S7H8_TETNG	Q4s7h8 tetraodon n
31	1513	54.4	317	2	Q6ZMM9_HUMAN	Q6zmm9 homo sapien
32	1338	48.1	355	2	Q4TEC7_TETNG	Q4tec7 tetraodon n
33	1308.5	47.1	441	2	Q5WV19_LEGPL	Q5wv19 legionella
34	1307.5	47.0	441	2	Q5X3N1_LEGPA	Q5x3n1 legionella
35	1307.5	47.0	441	2	Q5ZTY7_LEGPH	Q5zty7 legionella
36	1273.5	45.8	429	1	SAHH_COXBU	Q83a77 coxiella bu
37	1269.5	45.6	436	1	SAHH_LEPIC	Q75fu8 leptospira
38	1269.5	45.6	436	1	SAHH_LEPIN	Q8exv1 leptospira
39	1222	43.9	466	1	SAHH_CHRVO	Q7nzf7 chromobacte
40	1192	42.9	438	2	Q5FRZ9_GLUOX	Q5frz9 gluconobact
41	1191	42.8	472	2	Q4LHV4_9BURK	Q4lhv4 burkholderi
42	1188	42.7	473	1	SAHH_BURMA	Q62g22 burkholderi
43	1188	42.7	473	1	SAHH_BURPS	Q63pt2 burkholderi
44	1187	42.7	475	1	SAHH_GEOSL	P61617 geobacter s
45	1185	42.6	478	1	SAHH_NITEU	Q82wl1 nitrosomona

ALIGNMENTS

RESULT 1
Q96PK4_HUMAN
ID Q96PK4_HUMAN PRELIMINARY; PRT; 530 AA.
AC Q96PK4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE S-adenosylhomocysteine hydrolase-like protein (S-adenosylhomocysteine
DE hydrolase-like 1).
DE Adenosylhomocysteinase (EC 3.3.1.1) (AdoHcyase) (S-adenosyl-L-
DE homocysteine hydrolase).
GN Name=AHCYL1; ORFNames=RP11-180N18.1-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21901265; PubMed=11904675; DOI=10.1007/s00251-001-0402-z;
RA Dekker J.W., Budhia S., Angel N.Z., Cooper B.J., Clark G.J.,
RA Hart D.N., Kato M.;
RT "Identification of an S-adenosylhomocysteine hydrolase-like transcript
RT induced during dendritic cell differentiation.";
RL Immunogenetics 53:993-1001(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Martin S.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC homocysteine + adenosine.
CC -!- COFACTOR: NAD (By similarity).
CC -!- PATHWAY: Activated methyl cycle.
CC -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR EMBL; AF315687; AAL26869.1; -; mRNA.
DR EMBL; AL772411; CAH70965.1; -; Genomic_DNA.
DR HSSP; P10760; 1B3R.
DR GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR InterPro; IPR000043; Ad_hcy_hydrolase.
DR Pfam; PF05221; AdoHcyase; 1.
DR Pfam; PF00670; AdoHcyase_NAD; 1.
DR TIGRFAMs; TIGR00936; ahcY; 1.
DR PROSITE; PS00738; ADOHCYASE_1; 1.
DR PROSITE; PS00739; ADOHCYASE_2; 1.
KW Hydrolase; NAD; One-carbon metabolism.
SQ SEQUENCE 530 AA; 58951 MW; 974D23361A245D04 CRC64;

Query Match 100.0%; Score 2781; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 2.6e-208;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 2
Q80SW1_MOUSE
ID   Q80SW1_MOUSE PRELIMINARY;      PRT;   530 AA.
AC   Q80SW1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   S-adenosylhomocysteine hydrolase-like 1 (IP3R binding protein released
DE   with inositol 1,4,5-trisphosphate) (S-adenosylhomocysteine hydrolase-
DE   like protein).
GN   Name=Ahcyl1; Synonyms=Irbit;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Czech II;
RC   TISSUE=Mammary tumor metastatized to lung. MMTV-LTR/Wnt1 model.
RC   Expression driven by an MMTV-LTR enhancer.;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Czech II;
RC   TISSUE=Mammary tumor metastatized to lung. MMTV-LTR/Wnt1 model.
RC   Expression driven by an MMTV-LTR enhancer.;
RG   NIH MGC Project;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Cerebellum;
RX   MEDLINE=22526701; PubMed=12525476; DOI=10.1074/jbc.M210119200;
RA   Ando H., Mizutani A., Matsu-ura T., Mikoshiba K.;
RT   "IRBIT, a novel inositol 1,4,5-trisphosphate (IP3) receptor-binding
RT   protein, is released from the IP3 receptor upon IP3 binding to the
RT   receptor.";
RL   J. Biol. Chem. 278:10602-10612(2003).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21901265; PubMed=11904675; DOI=10.1007/s00251-001-0402-z;
RA   Dekker J.W., Budhia S., Angel N.Z., Cooper B.J., Clark G.J.,
RA   Hart D.N., Kato M.;
RT   "Identification of an S-adenosylhomocysteine hydrolase-like transcript
RT   induced during dendritic cell differentiation.";
RL   Immunogenetics 53:993-1001(2002).
CC   -!- MISCELLANEOUS: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ third party annotation (TPA) entry.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; BC018218; AAH18218.2; -; mRNA.
DR   EMBL; AB092504; BAC65166.1; -; mRNA.
DR   EMBL; BK000547; DAA00059.1; -; mRNA.
DR   HSSP; P10760; 1B3R.
DR   Ensembl; ENSMUSG00000027893; Mus musculus.
DR   MGI; MGI:2385184; Ahcyl1.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
SQ   SEQUENCE   530 AA;  58951 MW;  974D23361A245D04 CRC64;

  Query Match             100.0%;  Score 2781;  DB 2;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-208;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
```

```
Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
```

RESULT 3
Q502W8_HUMAN
ID Q502W8_HUMAN PRELIMINARY; PRT; 607 AA.
AC Q502W8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC   homocysteine + adenosine.
CC -!- COFACTOR: NAD (By similarity).
CC -!- PATHWAY: Activated methyl cycle.
CC -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR EMBL; BC095476; AAH95476.1; -; mRNA.
DR InterPro; IPR000043; Ad_hcy_hydrolase.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR011993; PH_type.
DR Pfam; PF05221; AdoHcyase; 1.
DR Pfam; PF00670; AdoHcyase_NAD; 1.
DR TIGRFAMs; TIGR00936; ahcY; 1.
DR PROSITE; PS00738; ADOHCYASE_1; 1.
DR PROSITE; PS00739; ADOHCYASE_2; UNKNOWN_1.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Hydrolase; Hypothetical protein; NAD; One-carbon metabolism.
FT NON_TER 1 1
SQ SEQUENCE 607 AA; 66683 MW; 349412833E2D5911 CRC64;

```
  Query Match            100.0%;  Score 2781;  DB 2;  Length 607;
  Best Local Similarity  100.0%;  Pred. No. 3.1e-208;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 137

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 197

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 257

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 317

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 377

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 437

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 497

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 557

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 607
```

RESULT 4
Q6P171_HUMAN
ID Q6P171_HUMAN PRELIMINARY; PRT; 623 AA.
AC Q6P171;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; BC065254; AAH65254.1; -; mRNA.
DR   HSSP; P10760; 1B3R.
DR   Ensembl; ENSG00000168710; Homo sapiens.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Hydrolase; Hypothetical protein; NAD; One-carbon metabolism.
FT   NON_TER      1      1
SQ   SEQUENCE   623 AA;  68648 MW;  B0A40ED58BAF3F23 CRC64;

  Query Match           100.0%;  Score 2781;  DB 2;  Length 623;
  Best Local Similarity  100.0%;  Pred. No. 3.2e-208;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 153

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 213

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 273

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 333

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 393

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 453

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 513

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 573

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 623


RESULT 5
SAHH2_HUMAN
ID   SAHH2_HUMAN    STANDARD;      PRT;   500 AA.
AC   O43865;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Putative adenosylhomocysteinase 2 (EC 3.3.1.1) (S-adenosyl-L-
DE   homocysteine hydrolase) (AdoHcyase).
GN   Name=AHCYL1; Synonyms=XPVKONA;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Skin;
RA   Volpe J.P.G., McDowell M., Jostes R.F., Afzal V., Sadinski W.,
RA   Trask B.J., Legerski R., Cleaver J.E.;
RT   "Complementation of chromosomal instability in the xeroderma
RT   pigmentosum variant by a gene on human chromosome 1 with homology to
RT   S-adenosyl homocysteine hydrolase.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Colon, and Placenta;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: Binds 1 NAD per subunit (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; U82761; AAC01960.1; -; mRNA.
DR   EMBL; BC007576; AAH07576.1; -; mRNA.
```

```
DR   EMBL; BC010681; AAH10681.1; -; mRNA.
DR   EMBL; BC016942; AAH16942.1; -; mRNA.
DR   HSSP; P10760; 1B3R.
DR   Ensembl; ENSG00000168710; Homo sapiens.
DR   HGNC; HGNC:344; AHCYL1.
DR   MIM; 607826; -.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   PANTHER; PTHR11784; Ad_hcy_hydrolase; 1.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
FT   REGION      251    418       NAD binding (By similarity).
FT   BINDING     125    125       Substrate (By similarity).
FT   BINDING     199    199       Substrate (By similarity).
FT   BINDING     224    224       Substrate (By similarity).
FT   BINDING     254    254       Substrate (By similarity).
FT   BINDING     258    258       Substrate (By similarity).
SQ   SEQUENCE   500 AA;  55628 MW;  FDEABFC70044B9DA CRC64;

  Query Match             94.5%;  Score 2627;  DB 1;  Length 500;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-196;
  Matches  500;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         31 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 60

Qy         91 PREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVG 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVG 120

Qy        151 CTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWW 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWW 180

Qy        211 CIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGK 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGK 240

Qy        271 LCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAAL 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAAL 300

Qy        331 KALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKN 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKN 360

Qy        391 SCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTV 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTV 420

Qy        451 PTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDD 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 PTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDD 480

Qy        511 QAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||||||
Db        481 QAKYLGLNKNGPFKPNYYRY 500


RESULT 6
Q9UG84_HUMAN
ID   Q9UG84_HUMAN PRELIMINARY;      PRT;   484 AA.
AC   Q9UG84;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Adenosylhomocysteinase (EC 3.3.1.1) (AdoHcyase) (S-adenosyl-L-
DE   homocysteine hydrolase) (Fragment).
GN   Name=DKFZp564A1523;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; AL049954; CAB43223.2; -; mRNA.
DR   PIR; T08681; T08681.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; Hypothetical protein; NAD; One-carbon metabolism.
FT   NON_TER       1      1
SQ   SEQUENCE   484 AA;  53868 MW;  8A55EF564F6B7E1D CRC64;

  Query Match             91.6%;  Score 2548;  DB 2;  Length 484;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-190;
  Matches  484;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         47 DMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFC 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFC 60

Qy        107 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLC 120

Qy        167 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMI 226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMI 180

Qy        227 LDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQK 286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQK 240

Qy        287 FDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPIC 346
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 FDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPIC 300

Qy        347 ALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDV 406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDV 360

Qy        407 TSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALAL 466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALAL 420

Qy        467 IELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPN 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPN 480

Qy        527 YYRY 530
              ||||
Db        481 YYRY 484
```

```
RESULT 7
Q5VSM0_HUMAN
ID   Q5VSM0_HUMAN PRELIMINARY;      PRT;   483 AA.
AC   Q5VSM0;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Adenosylhomocysteinase (EC 3.3.1.1) (AdoHcyase) (S-adenosyl-L-
DE   homocysteine hydrolase).
GN   Name=AHCYL1; ORFNames=RP11-180N18.1-002;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Martin S.;
RL   Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; AL772411; CAH70966.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000168710; Homo sapiens.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
SQ   SEQUENCE   483 AA;  53753 MW;  18DFC7E74697E1D4 CRC64;

  Query Match             91.4%;  Score 2542;  DB 2;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 1e-189;
  Matches  483;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         48 MQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCV 107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCV 60

Qy        108 KNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCA 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCA 120

Qy        168 LGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMIL 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMIL 180

Qy        228 DDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKF 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKF 240

Qy        288 DNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICA 347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICA 300

Qy        348 LQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVT 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVT 360

Qy        408 SLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALI 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALI 420

Qy        468 ELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNY 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNY 480

Qy        528 YRY 530
              |||
Db        481 YRY 483


RESULT 8
Q5F3Q3_CHICK
ID   Q5F3Q3_CHICK PRELIMINARY;      PRT;   540 AA.
AC   Q5F3Q3;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=RCJMB04_10b17;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CB; TISSUE=Bursa;
RA   Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA   Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA   Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT   "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT   genefunction analysis.";
RL   Genome Biol. 6:R6-R6(2005).
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; AJ851597; CAH65231.1; -; mRNA.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   540 AA;  59231 MW;  66E479BFA2368F92 CRC64;

  Query Match             88.8%;  Score 2469;  DB 2;  Length 540;
  Best Local Similarity   95.7%;  Pred. No. 6.2e-184;
  Matches  473;  Conservative    8;  Mismatches    9;  Indels    4;  Gaps    1;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||:|:    |  ||| ||||| |||||:||||||||||||||||||||||||||||||
Db          1 MSVPE----PAGGEEAKQAKEAEDAEKFSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 56

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 116

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 176

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        177 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNVDGWQANMILDDGGDLTHWVYKK 236

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 YPNIFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 296

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              |||||||||||||| ||||||||||||||||||||||:||||||||||||||||||||||
Db        297 LKRTTDVMFGGKQVAVCGYGEVGKGCCAALKALGAIVYVTEIDPICALQACMDGFRVVKL 356

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 SEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 416
```

```
Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 476

Qy        481 VYLLPKKMDEYVAS 494
              ||||||||    :|
Db        477 VYLLPKKMGSRFSS 490


RESULT 9
Q7T2B5_BRARE
ID   Q7T2B5_BRARE PRELIMINARY;      PRT;   512 AA.
AC   Q7T2B5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   S-adenosylhomocysteine hydrolase-like 1.
GN   Name=ahcyl1;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC054614; AAH54614.1; -; mRNA.
DR   HSSP; P10760; 1B3R.
DR   Ensembl; ENSDARG00000001500; Danio rerio.
DR   ZFIN; ZDB-GENE-021206-7; ahcyl1.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase.
SQ   SEQUENCE   512 AA;  56775 MW;  3E0A09DFFE086683 CRC64;

  Query Match             88.7%;  Score 2466;  DB 2;  Length 512;
  Best Local Similarity   92.1%;  Pred. No. 9.9e-184;
  Matches  464;  Conservative   22;  Mismatches   16;  Indels    2;  Gaps    1;

Qy         29 SFMATVTKAPKKQIQFA--DDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD 86
              :| |  |     |:||||  :| |||:||||||||||||||||||||||||||||||||||
Db          9 APTAVPTHKSVKKIQFANQEDKQEFSKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD 68

Qy         87 DEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGA 146
              || ||||| | ||||||:|||||||||||||||||||||||||||||||||||||||||||
Db         69 DETSPREKSQVNSKGSSDFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGA 128

Qy        147 KIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESED 206
              |: |||||||||||||||| ||||||||:|||||||||||||||||:||:||||||||||
Db        129 KVAGCTHITAQTAVLIETLVALGAQCRWTACNIYSTQNEVAAALAESGVSVFAWKGESED 188

Qy        207 DFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLS 266
              ||||||||||||||||||||||||||||||||||||:||||:|||||||||||||||||||
Db        189 DFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPSVFKKVRGIVEESVTGVHRLYQLS 248

Qy        267 KAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGC 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 KAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGC 308

Qy        327 CAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLD 386
              |:|||||||| :|||||||||||||||||||||:||:|:|||||||||||||||||||||
Db        309 CSALKALGAIVCVTEIDPICALQACMDGFRVVKLNEVVRQMDMVITCTGNKNVVTREHLD 368

Qy        387 RMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLS 446
              |||| |||||||||||||| ||||||||||||||||||||||||||:|||||||||||||
Db        369 RMKNGCIVCNMGHSNTEIDVASLRTPELTWERVRSQVDHVIWPDGKRVILLAEGRLLNLS 428

Qy        447 CSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTE 506
              |||||:|||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        429 CSTVPSFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPNFDAHLTE 488

Qy        507 LTDDQAKYLGLNKNGPFKPNYYRY 530
              |:|:||||:|||||||||||||||
Db        489 LSDEQAKYMGLNKNGPFKPNYYRY 512


RESULT 10
Q66IP1_XENLA
ID   Q66IP1_XENLA PRELIMINARY;      PRT;   520 AA.
AC   Q66IP1;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC86404 protein.
GN   Name=MGC86404;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```

```
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RA   Klein S., Gerhard D.S.;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC081269; AAH81269.1; -; mRNA.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
SQ   SEQUENCE   520 AA;  57839 MW;  C73583CF180234C3 CRC64;

  Query Match             88.4%;  Score 2459;  DB 2;  Length 520;
  Best Local Similarity   87.4%;  Pred. No. 3.6e-183;
  Matches  464;  Conservative   32;  Mismatches   23;  Indels   12;  Gaps    3;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||:|:|:           |   |||: ||| |:| ||| ||||||||||:|| |||||
Db          1 MSLPEAV----------TTDEPAAAEKFLFMAPVSKTPKK-IQFADDMQEFSKFRTKTGR 49

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              |||||||||||||||| ||||||||||:|||:|||  |||| :||||||||||:||||||
Db         50 RSLSRSISQSSTDSYSSVASYTDSSDDEMSPRDKQQITSKGSGDFCVKNIKQADFGRREI 109

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              :||||:| ||||||||||||||| ||||||||||||| ||||||||||||||||||||:
Db        110 DIAEQEMMALISLRKRAQGEKPLLGAKIVGCTHITAQTGVLIETLCALGAQCRWSACNIF 169

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMD-GWQANMILDDGGDLTHWVYK 239
              |||||||||||:|||:|||||||||||||||||||| | ||||||||||||||||||||
Db        170 STQNEVAAALAESGVAIFAWKGESEDDFWWCIDRCVNADSGWQANMILDDGGDLTHWVYK 229

Qy        240 KYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILD 299
              |||:|: :|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 KYPHVYTQIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILD 289

Qy        300 GLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVK 359
              ||||:|||||||||||||||||||||||||||:|:|::|||||||||||||:||||||
Db        290 GLKRSTDVMFGGKQVVVCGYGEVGKGCCAALKALGAVVHVTEIDPICALQACMEGFRVVK 349

Qy        360 LNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERV 419
              |:|:|||||||||||||| ||||||||| |||||||||||||||| ||||||||:
Db        350 LSEIIRQVDVVITCTGNKNVVNREHLDRMKNGCIVCNMGHSNTEIDVTSLHTPELTWERI 409

Qy        420 RSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQ 479
              ||||||:|||||||:|||||||||||||||:|||| ||||||:||||||||||||||||
Db        410 RSQVDHIIWPDGKRIVLLAEGRLLNLSCSTIPTFVHSITATTEALALIELYNAPEGRYKQ 469

Qy        480 DVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              |||||||||||||||||:|||||||||:|||||||||||||||||||||||
Db        470 DVYLLPKKMDEYVASLHLPTFDAHLTELTEDQAKYLGLNKNGPFKPNYYRY 520


RESULT 11
SAHH3_HUMAN
ID   SAHH3_HUMAN    STANDARD;      PRT;   611 AA.
AC   Q96HN2; O94917;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Putative adenosylhomocysteinase 3 (EC 3.3.1.1) (S-adenosyl-L-
DE   homocysteine hydrolase) (AdoHcyase).
GN   Name=KIAA0828;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Brain;
RX   MEDLINE=99156230; PubMed=10048485;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 5:355-364(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Lung, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: Binds 1 NAD per subunit (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AB020635; BAA74851.1; ALT_INIT; mRNA.
DR   EMBL; BC008349; AAH08349.1; -; mRNA.
DR   EMBL; BC024325; AAH24325.1; -; mRNA.
DR   HSSP; P23526; 1LI4.
DR   Ensembl; ENSG00000158467; Homo sapiens.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   PANTHER; PTHR11784; Ad_hcy_hydrolase; 1.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
FT   REGION      362    529       NAD binding (By similarity).
FT   BINDING     236    236       Substrate (By similarity).
FT   BINDING     310    310       Substrate (By similarity).
```

```
FT   BINDING     335    335       Substrate (By similarity).
FT   BINDING     365    365       Substrate (By similarity).
FT   BINDING     369    369       Substrate (By similarity).
FT   CONFLICT    122    122       Missing (in Ref. 1).
SQ   SEQUENCE   611 AA;  66721 MW;  A76B1275FC0D891D CRC64;

  Query Match             88.3%;  Score 2456;  DB 1;  Length 611;
  Best Local Similarity   84.9%;  Pred. No. 7.7e-183;
  Matches  468;  Conservative   31;  Mismatches   24;  Indels   28;  Gaps    3;

Qy          8 PLPGVGE-----------ELKQAKEIEDAEKY-------------SFMATVTKAP---K 39
              |:|| |              :: ||  :: :: :             |   |||:||   |
Db         61 PVPGPGSGPAAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 120

Qy         40 KQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNS 99
              |||||||  ||| | ||| ||||||||||||||||||||||||||||||| |||:||| ||
Db        121 KQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNS 180

Qy        100 KGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTA 159
              ||||:|||||||||||||||||:| ||::|||||||||||||||||||||||||||||
Db        181 KGSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGAKIVGCTHITAQTA 240

Qy        160 VLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMD 219
              ||:||| ||||||||:||||||| |||||||||:|  ||||||||||||||||||||::
Db        241 VLMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGESEDDFWWCIDRCVNVE 300

Qy        220 GWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVN 279
              ||| |||||||||||||:|||||:||||:|||||||||||||||||||||||||||||||
Db        301 GWQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLSKAGKLCVPAMNVN 360

Qy        280 DSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYI 339
              ||||||||||||||||||||||||:|||||||||||||||||||||||||:|:|||:
Db        361 DSVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGCCAALKAMGSIVYV 420

Qy        340 TEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGH 399
              |||||||||||||||:||||||||||:|||||||||||||||||||||||||||||||||
Db        421 TEIDPICALQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLDRMKNSCIVCNMGH 480

Qy        400 SNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITA 459
              ||||||| |||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        481 SNTEIDVASLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLNLSCSTVPTFVLSITA 540

Qy        460 TTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNK 519
              ||||||||||||||||||||||||||||||||||||:||||||||||:|||||||||
Db        541 TTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTELTDEQAKYLGLNK 600

Qy        520 NGPFKPNYYRY 530
              |||||||||||
Db        601 NGPFKPNYYRY 611


RESULT 12
Q68FL4_MOUSE
ID   Q68FL4_MOUSE PRELIMINARY;      PRT;   613 AA.
AC   Q68FL4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   RIKEN cDNA 4631427C17.
GN   Name=4631427C17Rik;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Director MGC Project;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; BC079660; AAH79660.1; -; mRNA.
DR   Ensembl; ENSMUSG00000029772; Mus musculus.
DR   MGI; MGI:1921590; 4631427C17Rik.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
SQ   SEQUENCE   613 AA;  66899 MW;  4413DBF1F4825860 CRC64;

  Query Match             88.2%;  Score 2454;  DB 2;  Length 613;
  Best Local Similarity   84.9%;  Pred. No. 1.1e-182;
  Matches  468;  Conservative   30;  Mismatches   25;  Indels   28;  Gaps    3;

Qy          8 PLPGVGE-----------ELKQAKEIEDAEKY-------------SFMATVTKAP---K 39
              | || |              :: ||  :: :: :             |   |||:||   |
Db         63 PAPGPGSGPTAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 122

Qy         40 KQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNS 99
              |||||||  ||| | ||| ||||||||||||||||||||||||||||||| |||:||| ||
Db        123 KQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNS 182

Qy        100 KGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTA 159
              ||||:|||||||||||||||||:| ||::|||||||||||||||||||||||||||||||
Db        183 KGSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGAKIVGCTHITAQTA 242

Qy        160 VLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMD 219
              ||:||| ||||||||:||||||| |||||||||:|  ||||||||||||||||||||::
Db        243 VLMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGESEDDFWWCIDRCVNVE 302

Qy        220 GWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVN 279
              ||| |||||||||||||:|||||:||||:|||||||||||||||||||||||||||||||
Db        303 GWQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLSKAGKLCVPAMNVN 362

Qy        280 DSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYI 339
              ||||||||||||||||||||||||:|||||||||||||||||||||||||:|:|||:
Db        363 DSVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGCCAALKAMGSIVYV 422

Qy        340 TEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGH 399
              |||||||||||||||:||||||||||:|||||||||||||||||||||||||||||||||
Db        423 TEIDPICALQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLDRMKNSCIVCNMGH 482
```

```
Qy        400 SNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITA 459
              ||||||| |||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        483 SNTEIDVASLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLNLSCSTVPTFVLSITA 542

Qy        460 TTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNK 519
              ||||||||||||||||||||||||||||||||||||||||:|||||||||:|||||||||
Db        543 TTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTELTDEQAKYLGLNK 602

Qy        520 NGPPFKPNYYRY 530
              ||||||||||||
Db        603 NGPPFKPNYYRY 613


RESULT 13
Q5BL43_XENTR
ID   Q5BL43_XENTR PRELIMINARY;      PRT;   588 AA.
AC   Q5BL43;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein LOC549163.
GN   Name=LOC549163;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Silurana.
OX   NCBI_TaxID=8364;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RA   Klein S., Gerhard D.S.;
RL   Submitted (FEB-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC090609; AAH90609.1; -; mRNA.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   InterPro; IPR007698; AlaDh_PNT_C.
DR   InterPro; IPR001379; Egg_lysin.
DR   InterPro; IPR001093; IMPDh/GMPRtase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; UNKNOWN_1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   588 AA;  64721 MW;  2CBAD9A46134581C CRC64;

  Query Match             88.0%;  Score 2446.5;  DB 2;  Length 588;
  Best Local Similarity   85.6%;  Pred. No. 4e-182;
  Matches  465;  Conservative   33;  Mismatches   28;  Indels   17;  Gaps    3;

Qy          4 PDAMPLPGVGEELKQA--------------KEIEDAEKYSFMATVTKAPK--KQIQFADD 47
              ||  |:| : : :| |               |:  | |  |    |::||:  |:|||||
Db         47 PDNSPIPALSQPIKGAPASAMKRSDPTQRPKDGGD-EAGSPEGAVSEAPRTVKKIQFADQ 105

Qy         48 MQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCV 107
               :|| | ||| |||||||||||||||||||||||||||||| |||:||| ||||||:|||
Db        106 KREFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSKGSSDFCV 165

Qy        108 KNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCA 167
              |||||||||||||||||:| ||::|||||||||||||||||||||||||||||||||| |
Db        166 KNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLGA 225

Qy        168 LGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMIL 227
              |||||||:||||||| |||||||||:|| ||:|||||||||||||||||::||| ||||
Db        226 LGAQCRWAACNIYSTLNEVAAALAESGVPVFSWKGESEDDFWWCIDRCVNVEGWQPNMIL 285

Qy        228 DDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKF 287
              |||||||||:|||||:||||:||||||||||||||||||||||||||||||||||||||
Db        286 DDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKF 345

Qy        288 DNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICA 347
              |||||||||||||||||:|||||||||||||||||||||||:|:|||:||||||||
Db        346 DNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGCCAALKAMGSIVYVTEIDPICA 405

Qy        348 LQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVT 407
              |||||||||:|||||||||:|||||||||||||||||| ||||||||||||||||||||
Db        406 LQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLDRMKNGCIVCNMGHSNTEIDVA 465

Qy        408 SLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALI 467
              |||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        466 SLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLNLSCSTVPTFVLSITATTQALALI 525

Qy        468 ELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNY 527
              |||||||||||||||||||||||||||||:|||||||:|:|||||||||||||||||||
Db        526 ELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTELSDEQAKYLGLNKNGPFKPNY 585

Qy        528 YRY 530
              |||
Db        586 YRY 588


RESULT 14
Q8BIH1_MOUSE
ID   Q8BIH1_MOUSE PRELIMINARY;      PRT;   508 AA.
AC   Q8BIH1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
DE   library, clone:E130105J10 product:SIMILAR TO S-ADENOSYLHOMOCYSTEINE
DE   HYDROLASE-LIKE 1 homolog.
GN   Name=4631427C17Rik;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
```

```
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; AK053527; BAC35415.1; -; mRNA.
DR   HSSP; P10760; 1D4G.
DR   Ensembl; ENSMUSG00000029772; Mus musculus.
DR   MGI; MGI:1921590; 4631427C17Rik.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
SQ   SEQUENCE   508 AA;  56743 MW;  25CC387F93AA6E39 CRC64;

  Query Match           87.7%;  Score 2439;  DB 2;  Length 508;
  Best Local Similarity   90.3%;  Pred. No. 1.3e-181;
  Matches  457;  Conservative   24;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         25 AEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDS 84
              : |  ::        | ||||||  ||| | ||| ||||||||||||||||||||||||||
Db          3 SSKKKYIVNSNSGIKAQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDS 62

Qy         85 SDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLA 144
              |||| |||:||| |||||||:||||||||||||||||||:| ||::||||||||||||
Db         63 SDDETSPRDKQQKNSKGSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLA 122

Qy        145 GAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGES 204
              |||||||||||||||:||| |||||||:||||||| |||||||||:|  ||||||||
Db        123 GAKIVGCTHITAQTAVLMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGES 182

Qy        205 EDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQ 264
              |||||||||||::||| |||||||||||:||||||:||||:|||||||||||||||||
Db        183 EDDFWWCIDRCVNVEGWQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQ 242

Qy        265 LSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGK 324
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        243 LSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGK 302

Qy        325 GCCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREH 384
              ||||||||:|:|||:|||||||||||||||||||||||||||||||:|||||||||||||
Db        303 GCCAALKAMGSIVYVTEIDPICALQACMDGFRVVKLNEVIRQVDIVITCTGNKNVVTREH 362

Qy        385 LDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLN 444
              |||||||||||||||||| ||||||||||||||||||||||||:||||||||||
Db        363 LDRMKNSCIVCNMGHSNTEIDVASLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLN 422

Qy        445 LSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHL 504
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        423 LSCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHL 482

Qy        505 TELTDDQAKYLGLNKNGPFKPNYYRY 530
              |||||:||||||||||||||||||||
Db        483 TELTDEQAKYLGLNKNGPFKPNYYRY 508


RESULT 15
Q68EX1_XENLA
ID   Q68EX1_XENLA PRELIMINARY;      PRT;   588 AA.
AC   Q68EX1;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC84148 protein.
GN   Name=MGC84148;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
```

```
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RA   Klein S., Gerhard D.S.;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC080079; AAH80079.1; -; mRNA.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
SQ   SEQUENCE   588 AA;  64853 MW;  81A3869C48D026B8 CRC64;

  Query Match             87.7%;  Score 2438.5;  DB 2;  Length 588;
  Best Local Similarity   85.6%;  Pred. No. 1.7e-181;
  Matches  465;  Conservative   32;  Mismatches   29;  Indels   17;  Gaps    3;

Qy          4 PDAMPLPGVGEELKQA--------------KEIEDAEKYSFMATVTKAPK--KQIQFADD 47
              |:  |:| :   :| |              |:  | |  |  |::||:  |:|||||
Db         47 PEISPIPALSLPMKGAPASAMKRTDPTLRPKDGGD-EAGSPEGAVSEAPRTVKKIQFADQ 105

Qy         48 MQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCV 107
                :|| | ||| |||||||||||||||||||||||||||||| |||:||| ||||||:|||
Db        106 KREFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSKGSSDFCV 165

Qy        108 KNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCA 167
              |||||||||||||||||:| ||::||||||||||||||||||||||||||||||||| |
Db        166 KNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLGA 225

Qy        168 LGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMIL 227
              |||||||:|||||||  |||||||||  |||||:||||||||||||||||::||| ||||
Db        226 LGAQCRWAACNIYSTLNEVAAALAENGVAVFSWKGESEDDFWWCIDRCVNVEGWQPNMIL 285

Qy        228 DDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKF 287
              |||||||||:||||||:||||:|||||||||||||||||||||||||||||||||||||
Db        286 DDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKF 345

Qy        288 DNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICA 347
              |||||||||||||||||:|||||||||||||||||||||||:|:|||:||||||||
Db        346 DNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGCCAALKAMGSIVYVTEIDPICA 405

Qy        348 LQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVT 407
              ||||||||:||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        406 LQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLDRMKNGCIVCNMGHSNTEIDVA 465

Qy        408 SLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALI 467
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        466 SLRTPELTWERVRSQVDHVIWPDGKRMVLLAEGRLLNLSCSTVPTFVLSITATTQALALI 525

Qy        468 ELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNY 527
              |||||||||||||||||||||||||||||:||||||||:|:|||||||||||||||||
Db        526 ELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTELSDEQAKYLGLNKNGPFKPNY 585

Qy        528 YRY 530
              |||
Db        586 YRY 588


Search completed: December 14, 2005, 08:20:12
Job time : 175.208 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:13:48 ; Search time 7.3817 Seconds
(without alignments)
1164.808 Million cell updates/sec

Title: US-10-683-610-1_COPY_1_104
Perfect score: 521
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....SDDEVSPREKQQTNSKGSSN 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	367	70.4	500	1	US-08-896-005-1	Sequence 1, Appli
2	86.5	16.6	414	1	US-07-667-276A-4	Sequence 4, Appli
3	81.5	15.6	542	2	US-09-538-092-289	Sequence 289, App
4	79.5	15.3	703	2	US-08-910-925-4	Sequence 4, Appli
5	79	15.2	998	2	US-09-949-016-7757	Sequence 7757, Ap
6	76	14.6	754	2	US-09-976-594-375	Sequence 375, App
7	76	14.6	760	2	US-09-949-016-11129	Sequence 11129, A
8	75.5	14.5	124	2	US-09-270-767-34893	Sequence 34893, A
9	75.5	14.5	124	2	US-09-270-767-50110	Sequence 50110, A
10	75.5	14.5	414	2	US-09-248-796A-19046	Sequence 19046, A
11	75	14.4	504	1	US-08-896-005-3	Sequence 3, Appli
12	74.5	14.3	442	2	US-09-248-796A-19805	Sequence 19805, A
13	74	14.2	191	2	US-09-248-796A-15111	Sequence 15111, A
14	72.5	13.9	208	2	US-09-925-637-52	Sequence 52, Appl
15	72.5	13.9	439	2	US-08-860-368B-14	Sequence 14, Appl
16	72.5	13.9	441	2	US-09-248-796A-26086	Sequence 26086, A
17	72.5	13.9	984	1	US-08-257-073-3	Sequence 3, Appli
18	72.5	13.9	984	1	US-08-184-009-120	Sequence 120, App
19	72.5	13.9	984	1	US-08-458-356-120	Sequence 120, App
20	72.5	13.9	984	2	US-08-460-736-120	Sequence 120, App
21	72.5	13.9	984	2	US-09-535-370-120	Sequence 120, App
22	72.5	13.9	984	2	US-09-663-667-120	Sequence 120, App
23	72	13.8	828	2	US-09-327-984A-2	Sequence 2, Appli
24	72	13.8	1850	2	US-09-620-093A-5	Sequence 5, Appli
25	72	13.8	3256	2	US-09-919-172-98	Sequence 98, Appl
26	72	13.8	3256	2	US-09-976-594-22	Sequence 22, Appl
27	72	13.8	3256	2	US-09-919-039-21	Sequence 21, Appl
28	71.5	13.7	129	2	US-09-248-796A-24096	Sequence 24096, A
29	71.5	13.7	233	2	US-09-270-767-45949	Sequence 45949, A
30	71.5	13.7	291	2	US-08-860-368B-17	Sequence 17, Appl
31	71.5	13.7	379	2	US-08-860-368B-15	Sequence 15, Appl
32	71.5	13.7	568	1	US-08-320-559-30	Sequence 30, Appl
33	71.5	13.7	568	2	US-08-545-860D-30	Sequence 30, Appl
34	71.5	13.7	568	2	US-09-538-092-1114	Sequence 1114, Ap
35	71.5	13.7	568	4	PCT-US94-04496-30	Sequence 30, Appl
36	71	13.6	265	2	US-09-248-796A-18812	Sequence 18812, A
37	70.5	13.5	236	2	US-09-248-796A-21575	Sequence 21575, A
38	70.5	13.5	684	2	US-09-252-991A-28604	Sequence 28604, A
39	70	13.4	380	2	US-09-270-767-43626	Sequence 43626, A
40	70	13.4	533	2	US-09-270-767-42903	Sequence 42903, A
41	70	13.4	2375	2	US-09-538-092-1131	Sequence 1131, Ap
42	69.5	13.3	249	2	US-09-248-796A-20310	Sequence 20310, A
43	69.5	13.3	268	2	US-08-961-083-140	Sequence 140, App
44	69.5	13.3	268	2	US-09-536-784-140	Sequence 140, App
45	69.5	13.3	268	2	US-09-765-271-140	Sequence 140, App

ALIGNMENTS

RESULT 1
US-08-896-005-1
; Sequence 1, Application US/08896005
; Patent No. 5854023
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,005
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0337 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADTUT04
; CLONE: 1519044
US-08-896-005-1

Query Match 70.4%; Score 367; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 7.3e-35;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          31 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 90
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 60

Qy          91 PREKQQTNSKGSSN 104
               ||||||||||||||
Db          61 PREKQQTNSKGSSN 74


RESULT 2
US-07-667-276A-4
; Sequence 4, Application US/07667276A
; Patent No. 5470971
;  GENERAL INFORMATION:
;    APPLICANT:  Kondo, Keiji
;    APPLICANT:  Inouye, Masayori
;    TITLE OF INVENTION:  STRESS-INDUCED PROTEINS, GENES CODING
;    TITLE OF INVENTION:  THEREFOR, TRANSFORMED CELLS OF ORGANISMS, METHODS AND
;    TITLE OF INVENTION:  APPLICATIONS
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weiser & Associates
;      STREET:  230 S. Fifteenth Street, Suite 500
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/667,276A
;      FILING DATE:  11-MAR-1991
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiser, Gerard J.
;      REGISTRATION NUMBER:  19,763
;      REFERENCE/DOCKET NUMBER:  377.5351P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-875-8383
;      TELEFAX:  215-875-8394
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  414 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-667-276A-4

  Query Match             16.6%;  Score 86.5;  DB 1;  Length 414;
  Best Local Similarity   31.4%;  Pred. No. 0.083;
  Matches   32;  Conservative   16;  Mismatches   39;  Indels   15;  Gaps    3;

Qy          17 KQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYS 76
               ||||| | |:  |  :: : :      ::   |     : :   | | | | ||:|| |
Db          18 KQAKE-EKAKAVSSSSSESSSSSSSSSESESESESESESSSSSSSSDSESSSSSSSDSES 76

Qy          77 ----------SAASYTDSSDDEVSPREKQQT---NSKGSSN 104
                       |::| :||| ||    ||::|    || ||:
Db          77 EAETKKEESKDSSSSSSDSSSDEEEEEEKEETKKEESKESSS 118


RESULT 3
US-09-538-092-289
; Sequence 289, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 289
;  LENGTH: 542
;  TYPE: PRT
;  ORGANISM: Saccharomyces cerevisiae
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (0)...(0)
;  OTHER INFORMATION: Polypeptide Accession Number YGL028C
US-09-538-092-289

  Query Match             15.6%;  Score 81.5;  DB 2;  Length 542;
  Best Local Similarity   26.7%;  Pred. No. 0.47;
  Matches   27;  Conservative   20;  Mismatches   45;  Indels    9;  Gaps    2;

Qy          11 GVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTK-------FPTKTGRRSL 63
               ||||   :      :: :|: |  |    :|   :   || |:         | :   |
Db         174 GVGEANVEGN--TQSQHHSYQAAATSTLNQQTSTSIASQESTESTNTPTSSSTSSSTSSS 231

Qy          64 SRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
               : | : ||| | :|::: : :|    | :|   | |:|||:
Db         232 TSSSTSSSTSSSTSSSTSSSTSSSTSSTQETAATTSEGSSS 272


RESULT 4
US-08-910-925-4
; Sequence 4, Application US/08910925
; Patent No. 6162601
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN PININ SPLICE VARIANT
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/910,925
;      FILING DATE:  Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0365 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  703 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1684843
US-08-910-925-4

  Query Match             15.3%;  Score 79.5;  DB 2;  Length 703;
  Best Local Similarity   27.2%;  Pred. No. 1.2;
  Matches   31;  Conservative   20;  Mismatches   52;  Indels   11;  Gaps    3;

Qy          1 MSMPDAMPLPGVGE---ELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTK 57
              :| |: :|| | :   :| | :     |:  |:    |  :   :    :   :|: ||
Db        512 LSQPETLPL-AVSQPPPQLIQRQGHLPPERKEFLVESVKLTEVPTEPVLTVHSESKYETK 570

Qy         58 TGRRS-------LSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |  ||        |:| |:||: | ||::| : ||    |         |  ||:
Db        571 TRSRSRGRARNRTSKSRSRSSSSSSSSSSSTSSSSGSSSSSGSSSSRTSSSSSS 624


RESULT 5
US-09-949-016-7757
; Sequence 7757, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7757
;   LENGTH: 998
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-7757

  Query Match             15.2%;  Score 79;  DB 2;  Length 998;
  Best Local Similarity   29.8%;  Pred. No. 2.1;
  Matches   31;  Conservative   17;  Mismatches   48;  Indels    8;  Gaps    4;

Qy          2 SMPDAMPLPGVGEEL-KQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ::|      |   ||  :: : ::: ||       :|  :    || :: |: |   | |
Db        481 TIPTTQAPPLSREEFYREQRRLKEEEKKK-----SKLDEFTNDFAKELMEYKKI-QKERR 534

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              || ||| |  |  || | :||| |     | | :  : |   |:
Db        535 RSFSRSKSPYSGSSY-SRSSYTYSKSRSGSTRSRSYSRSFSRSH 577


RESULT 6
US-09-976-594-375
; Sequence 375, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 375
;   LENGTH: 754
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 3688791CD1
US-09-976-594-375

  Query Match             14.6%;  Score 76;  DB 2;  Length 754;
  Best Local Similarity   29.9%;  Pred. No. 3.3;
  Matches   29;  Conservative   14;  Mismatches   34;  Indels   20;  Gaps    4;

Qy         13 GEELKQAKEIE----DAEKYSF-----MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSL 63
              |:|:   :|||    ||    |     ::     || :::          |   |  : |
Db        147 GQEV--VREIENQKTDAASKPFAEVRILSCGELIPKSKVK---------KEEKKRHKSSS 195

Qy         64 SRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSK 100
              | | | | :|| | : | :|||| | :  || :   |
Db        196 SSSSSSSDSDSSSDSQSSSDSSDSESATEEKSKKRKK 232


RESULT 7
US-09-949-016-11129
; Sequence 11129, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11129
;   LENGTH: 760
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-11129

  Query Match             14.6%;  Score 76;  DB 2;  Length 760;
  Best Local Similarity   29.9%;  Pred. No. 3.3;
  Matches   29;  Conservative   14;  Mismatches   34;  Indels   20;  Gaps    4;

Qy         13 GEELKQAKEIE----DAEKYSF-----MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSL 63
              |:|:   :|||    ||    |     ::     || :::          |   |  : |
Db        153 GQEV--VREIENQKTDAASKPFAEVRILSCGELIPKSKVK---------KEEKKRHKSSS 201

Qy         64 SRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSK 100
              | | | | :|| | : | :|||| | :  || :   |
Db        202 SSSSSSSDSDSSSDSQSSSDSSDSESATEEKSKKRKK 238


RESULT 8
US-09-270-767-34893
; Sequence 34893, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34893
```

```
;   LENGTH: 124
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-34893

  Query Match             14.5%;  Score 75.5;  DB 2;  Length 124;
  Best Local Similarity   26.2%;  Pred. No. 0.32;
  Matches   28;  Conservative   17;  Mismatches   37;  Indels   25;  Gaps    3;

Qy         23 EDAEKYSFMATVTK------------APKKQI----QFADDMQEFTKFPTKTGRRSLSRS 66
              ::  ||   :: ||            ||| |:    : |    :  :   || : | | |
Db          2 KEKRKYVHRSSRTKQLSKRMSQKEADAPKNQVEEHSEVASSEGKKVRRKRKTSKSSSSSS 61

Qy         67 ISQSSTDSYS---------SAASYTDSSDDEVSPREKQQTNSKGSSN 104
                |  |:|| |         |::| :|| |:| |  :   : |   |:
Db         62 SSSDSSDSSSESGSSSESDSSSSSSDSDDEEGSSSDSSGSGSDSDSS 108


RESULT 9
US-09-270-767-50110
; Sequence 50110, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50110
;   LENGTH: 124
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-50110

  Query Match             14.5%;  Score 75.5;  DB 2;  Length 124;
  Best Local Similarity   26.2%;  Pred. No. 0.32;
  Matches   28;  Conservative   17;  Mismatches   37;  Indels   25;  Gaps    3;

Qy         23 EDAEKYSFMATVTK------------APKKQI----QFADDMQEFTKFPTKTGRRSLSRS 66
              ::  ||   :: ||            ||| |:    : |    :  :   || : | | |
Db          2 KEKRKYVHRSSRTKQLSKRMSQKEADAPKNQVEEHSEVASSEGKKVRRKRKTSKSSSSSS 61

Qy         67 ISQSSTDSYS---------SAASYTDSSDDEVSPREKQQTNSKGSSN 104
                |  |:|| |         |::| :|| |:| |  :   : |   |:
Db         62 SSSDSSDSSSESGSSSESDSSSSSSDSDDEEGSSSDSSGSGSDSDSS 108


RESULT 10
US-09-248-796A-19046
; Sequence 19046, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19046
;   LENGTH: 414
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-19046

  Query Match             14.5%;  Score 75.5;  DB 2;  Length 414;
  Best Local Similarity   28.3%;  Pred. No. 1.7;
  Matches   30;  Conservative   16;  Mismatches   43;  Indels   17;  Gaps    4;

Qy          9 LPGVGEELKQ-AKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSI 67
              ||   ::|:    |:|: || |         |:     ::|    ::  |      | | |
Db         97 LPKFSKQLESIIDEVENQEKKS-------KPRNSSSDSEDYSSESESSTSDSESSSSDSD 149

Qy         68 S------QSSTDSYSSAASYTDSSDD---EVSPREKQQTNSKGSSN 104
              |        ||:|| ||::   || |:   |    ||   :|: | |
Db        150 SFSSDSESSSSDSESSSSDSEDSDDEEDKEDKEAEKDNKDSEDSEN 195


RESULT 11
US-08-896-005-3
; Sequence 3, Application US/08896005
; Patent No. 5854023
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/896,005
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0337 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  504 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  969078
US-08-896-005-3

  Query Match             14.4%;  Score 75;  DB 1;  Length 504;
  Best Local Similarity   45.2%;  Pred. No. 2.5;
  Matches   19;  Conservative    8;  Mismatches   13;  Indels    2;  Gaps    2;

Qy         64 SRSISQSSTDSYSSAASYTDSSDDE-VSPREKQQTNSKGSSN 104
              :|  |  || |:||   :| |||:| |||::  | || | ::
Db         27 ARRFSDVSTCSFSSTC-FTGSSDEEDVSPKDNHQRNSAGGTD 67


RESULT 12
US-09-248-796A-19805
; Sequence 19805, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
```

```
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19805
;   LENGTH: 442
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-19805

  Query Match             14.3%;  Score 74.5;  DB 2;  Length 442;
  Best Local Similarity   33.0%;  Pred. No. 2.4;
  Matches   31;  Conservative   11;  Mismatches   31;  Indels   21;  Gaps    6;

Qy         23 EDAEKYSFMATVTKAPKKQIQFADDMQEFTKFP----TKTGRRSLSRSIS------QSST 72
              :| | | |  :|:      |:  ||  ||   |     | | |:    |||       ||:
Db        193 DDVESY-FSGSVSTV--SSIEPLDD--EFVPPPQPPRTHTSRKRKHDSISPPASSDSSSS 247

Qy         73 DSY------SSAASYTDSSDDEVSPREKQQTNSK 100
               ||      ||::| | : |  |||  |::   |
Db        248 SSYVPQLIPSSSSSVTSNGDSPVSPTTKRKYTKK 281


RESULT 13
US-09-248-796A-15111
; Sequence 15111, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15111
;   LENGTH: 191
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-15111

  Query Match             14.2%;  Score 74;  DB 2;  Length 191;
  Best Local Similarity   23.4%;  Pred. No. 0.87;
  Matches   26;  Conservative   24;  Mismatches   45;  Indels   16;  Gaps    5;

Qy          2 SMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEF-----TKFPT 56
              |:||:   ||  :  | :          |     |:  :  : | |::::      || |:
Db         32 SLPDSK--PGTPKPSKTSGSNTGGGNDS-----TENKEDLLDFLDELEQHEKTKKTKKPS 84

Qy         57 KTGRRSLSRSISQSSTDSYSSAASYTDSSD---DEVSPREKQQTNSKGSSN 104
              |   :: |:  | |  |    :  : || ||   ::: |:: |:   : | |
Db         85 KLAPKTASKPTS-SDADDEPTTGTTTDKSDTTNEDIEPQQPQEKTEEESDN 134


RESULT 14
US-09-925-637-52
; Sequence 52, Application US/09925637
; Patent No. 6833253
; GENERAL INFORMATION:
;  APPLICANT: Choi
;  TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
;  FILE REFERENCE: PB560
;  CURRENT APPLICATION NUMBER: US/09/925,637
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/23773
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US 60/151,933
;  PRIOR FILING DATE: 1999-09-01
;  PRIOR APPLICATION NUMBER: US 08/781,986
;  PRIOR FILING DATE: 1997-01-03
;  PRIOR APPLICATION NUMBER: US 08/956,171
;  PRIOR FILING DATE: 1997-10-20
;  PRIOR APPLICATION NUMBER: US 60/009,861
;  PRIOR FILING DATE: 1996-01-06
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 208
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (174)..(174)
;   OTHER INFORMATION: Xaa equals any amino acid
US-09-925-637-52

  Query Match             13.9%;  Score 72.5;  DB 2;  Length 208;
  Best Local Similarity   27.6%;  Pred. No. 1.5;
  Matches   27;  Conservative   16;  Mismatches   38;  Indels   17;  Gaps    4;

Qy          5 DAMPLPGVGEEL-----KQAKEIEDAEKYSFMATVTKAPKKQIQFADD-MQEFTKFPTKT 58
              |  |:  ||:::      |: |  ||        |   || |  |: ||| ::|| |
Db         40 DEEPIKTVGKKIAELDEKKKKLTEDVNSKD-TAVRGKAVKDLIKNADDRLKEFEK----- 93

Qy         59 GRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQ 96
                        : |   :  | |: |: |::|  :| :|
Db         94 -----EEDAIKKSEQDFKKAKSHVDNIDNDVKRKEVKQ 126


RESULT 15
US-08-860-368B-14
; Sequence 14, Application US/08860368B
; Patent No. 6002072
;  GENERAL INFORMATION:
;    APPLICANT:  McMaster, Russell J.
;    APPLICANT:  Boeshore, Maury L.
;    APPLICANT:  Tricoli, David M.
;    APPLICANT:  Reynolds, John F.
;    APPLICANT:  Carney, Kim J.
;    APPLICANT:  Gonsalves, Dennis
;    TITLE OF INVENTION:  Papaya Ringspot Virus Coat Protein Gene
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rockey, Milnamow & Katz, Ltd.
;      STREET:  180 N. Stetson Avenue, 2 Prudential Plaza,
;      STREET:  Suite 4700
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  U.S.A.
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/860,368B
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mueller, Lisa L.
;      REGISTRATION NUMBER:  38,978
;      REFERENCE/DOCKET NUMBER:  SVS3801P0181US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-616-5400
;      TELEFAX:  312-616-5460
```

```
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  439 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-860-368B-14

  Query Match             13.9%;  Score 72.5;  DB 2;  Length 439;
  Best Local Similarity   28.7%;  Pred. No. 4.1;
  Matches   25;  Conservative   16;  Mismatches   35;  Indels   11;  Gaps    3;

Qy         11 GVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQS 70
              |: |:||: :: :: ||        |  :|:   | |  : :   |||| |    ::  |
Db        161 GLNEKLKEKEKQKEKEK-------EKQKEKEKDGASDGNDVST-STKTGERDRDVNVGTS 212

Qy         71 STDSYSSAASYTDSSDDEVSPREKQQT 97
              | :     |:||     | || | :|
Db        213 GTFTVPRIKSFTDKM---VLPRIKGKT 236


Search completed: December 14, 2005, 08:21:48
Job time : 8.3817 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:20:23 ; Search time 23.7855 Seconds
(without alignments)
1826.923 Million cell updates/sec

Title: US-10-683-610-1_COPY_1_104
Perfect score: 521
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....SDDEVSPREKQQTNSKGSSN 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	521	100.0	239	3	US-09-925-300-1379	Sequence 1379, Ap
2	521	100.0	530	5	US-10-876-285-2	Sequence 2, Appli
3	521	100.0	530	5	US-10-876-285-4	Sequence 4, Appli
4	521	100.0	530	5	US-10-876-285-6	Sequence 6, Appli
5	521	100.0	530	5	US-10-876-285-8	Sequence 8, Appli
6	521	100.0	530	5	US-10-683-610-1	Sequence 1, Appli
7	521	100.0	614	3	US-09-782-051-2	Sequence 2, Appli
8	288.5	55.4	455	5	US-10-450-763-42388	Sequence 42388, A
9	288.5	55.4	619	5	US-10-756-149-5569	Sequence 5569, Ap
10	286.5	55.0	257	4	US-10-424-599-212758	Sequence 212758,
11	273	52.4	564	5	US-10-467-903-1	Sequence 1, Appli
12	126.5	24.3	521	6	US-11-097-143-2781	Sequence 2781, Ap
13	86.5	16.6	414	6	US-11-004-418A-28	Sequence 28, Appl
14	84.5	16.2	321	4	US-10-424-599-169465	Sequence 169465,
15	83.5	16.0	2058	4	US-10-052-648A-20	Sequence 20, Appl
16	83	15.9	475	4	US-10-437-963-120425	Sequence 120425,
17	83	15.9	628	4	US-10-369-493-22576	Sequence 22576, A
18	83	15.9	2645	5	US-10-732-923-8232	Sequence 8232, Ap
19	82	15.7	2296	5	US-10-696-909A-46	Sequence 46, Appl
20	81.5	15.6	428	4	US-10-451-467A-448	Sequence 448, App
21	80.5	15.5	166	4	US-10-425-115-264771	Sequence 264771,
22	80.5	15.5	740	4	US-10-437-963-134499	Sequence 134499,
23	79.5	15.3	156	3	US-09-925-301-1154	Sequence 1154, Ap
24	79.5	15.3	176	6	US-11-097-143-40977	Sequence 40977, A
25	79.5	15.3	2752	5	US-10-696-909A-44	Sequence 44, Appl
26	79	15.2	948	4	US-10-308-448-15	Sequence 15, Appl
27	79	15.2	948	5	US-10-934-998-295	Sequence 295, App
28	79	15.2	1616	5	US-10-934-998-88	Sequence 88, Appl
29	78.5	15.1	275	4	US-10-424-599-231467	Sequence 231467,
30	78	15.0	1067	5	US-10-732-923-8348	Sequence 8348, Ap
31	77.5	14.9	537	6	US-11-097-143-36798	Sequence 36798, A
32	77.5	14.9	648	6	US-11-097-143-16119	Sequence 16119, A
33	77.5	14.9	776	4	US-10-437-963-134523	Sequence 134523,
34	77	14.8	406	4	US-10-451-467A-262	Sequence 262, App
35	76.5	14.7	149	4	US-10-425-114-40637	Sequence 40637, A
36	76.5	14.7	396	4	US-10-437-963-204429	Sequence 204429,
37	76.5	14.7	568	4	US-10-437-963-134407	Sequence 134407,
38	76.5	14.7	577	6	US-11-097-143-21738	Sequence 21738, A
39	76	14.6	244	4	US-10-424-599-271081	Sequence 271081,
40	76	14.6	333	6	US-11-097-143-33054	Sequence 33054, A
41	76	14.6	754	4	US-10-153-668-254	Sequence 254, App
42	76	14.6	913	4	US-10-425-114-54384	Sequence 54384, A
43	76	14.6	970	4	US-10-425-115-295229	Sequence 295229,
44	76	14.6	1406	6	US-11-097-143-42681	Sequence 42681, A
45	76	14.6	2421	6	US-11-097-143-23721	Sequence 23721, A

ALIGNMENTS

RESULT 1
US-09-925-300-1379
; Sequence 1379, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1379
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (229)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (231)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (234)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1379

Query Match 100.0%; Score 521; DB 3; Length 239;
Best Local Similarity 100.0%; Pred. No. 3.3e-45;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          86 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 145

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||||||||||||||||||||||||||||
Db        146 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 189
```

RESULT 2
US-10-876-285-2
; Sequence 2, Application US/10876285
; Publication No. US20050044582A1

```
; GENERAL INFORMATION:
;  APPLICANT: Kato, Masato
;  APPLICANT:  Angel, Nicola Zoe
;  APPLICANT:  Cooper, Benjamine Jay
;  APPLICANT:  Hart, Derek Nigel John
;  TITLE OF INVENTION: Novel Polynucleotides and uses therefor
;  FILE REFERENCE: 10338-18US
;  CURRENT APPLICATION NUMBER: US/10/876,285
;  CURRENT FILING DATE:  2004-06-24
;  PRIOR APPLICATION NUMBER: PR9741/01
;  PRIOR FILING DATE: 2001-12-24
;  PRIOR APPLICATION NUMBER: PCT/AU02/01761
;  PRIOR FILING DATE: 2002-12-24
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: human
US-10-876-285-2

  Query Match           100.0%;  Score 521;  DB 5;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 9e-45;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 3
US-10-876-285-4
; Sequence 4, Application US/10876285
; Publication No. US20050044582A1
; GENERAL INFORMATION:
;  APPLICANT: Kato, Masato
;  APPLICANT:  Angel, Nicola Zoe
;  APPLICANT:  Cooper, Benjamine Jay
;  APPLICANT:  Hart, Derek Nigel John
;  TITLE OF INVENTION: Novel Polynucleotides and uses therefor
;  FILE REFERENCE: 10338-18US
;  CURRENT APPLICATION NUMBER: US/10/876,285
;  CURRENT FILING DATE:  2004-06-24
;  PRIOR APPLICATION NUMBER: PR9741/01
;  PRIOR FILING DATE: 2001-12-24
;  PRIOR APPLICATION NUMBER: PCT/AU02/01761
;  PRIOR FILING DATE: 2002-12-24
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: Human
US-10-876-285-4

  Query Match           100.0%;  Score 521;  DB 5;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 9e-45;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 4
US-10-876-285-6
; Sequence 6, Application US/10876285
; Publication No. US20050044582A1
; GENERAL INFORMATION:
;  APPLICANT: Kato, Masato
;  APPLICANT:  Angel, Nicola Zoe
;  APPLICANT:  Cooper, Benjamine Jay
;  APPLICANT:  Hart, Derek Nigel John
;  TITLE OF INVENTION: Novel Polynucleotides and uses therefor
;  FILE REFERENCE: 10338-18US
;  CURRENT APPLICATION NUMBER: US/10/876,285
;  CURRENT FILING DATE:  2004-06-24
;  PRIOR APPLICATION NUMBER: PR9741/01
;  PRIOR FILING DATE: 2001-12-24
;  PRIOR APPLICATION NUMBER: PCT/AU02/01761
;  PRIOR FILING DATE: 2002-12-24
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: Mouse
US-10-876-285-6

  Query Match           100.0%;  Score 521;  DB 5;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 9e-45;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 5
US-10-876-285-8
; Sequence 8, Application US/10876285
; Publication No. US20050044582A1
; GENERAL INFORMATION:
;  APPLICANT: Kato, Masato
;  APPLICANT:  Angel, Nicola Zoe
;  APPLICANT:  Cooper, Benjamine Jay
;  APPLICANT:  Hart, Derek Nigel John
;  TITLE OF INVENTION: Novel Polynucleotides and uses therefor
;  FILE REFERENCE: 10338-18US
;  CURRENT APPLICATION NUMBER: US/10/876,285
;  CURRENT FILING DATE:  2004-06-24
;  PRIOR APPLICATION NUMBER: PR9741/01
;  PRIOR FILING DATE: 2001-12-24
;  PRIOR APPLICATION NUMBER: PCT/AU02/01761
;  PRIOR FILING DATE: 2002-12-24
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: Mouse
US-10-876-285-8

  Query Match           100.0%;  Score 521;  DB 5;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 9e-45;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
```

```
RESULT 6
US-10-683-610-1
; Sequence 1, Application US/10683610
; Publication No. US20050181450A1
; GENERAL INFORMATION:
;  APPLICANT: RIKEN
;  TITLE OF INVENTION: A novel inositol 1,4,5-trisphosphate (IP3) receptor binding prote
;  TITLE OF INVENTION:  IP3
;  TITLE OF INVENTION:  indicator
;  FILE REFERENCE: PH-1852
;  CURRENT APPLICATION NUMBER: US/10/683,610
;  CURRENT FILING DATE:  2003-10-10
;  PRIOR APPLICATION NUMBER: JP 2002-299429
;  PRIOR FILING DATE: 2002-10-11
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-683-610-1

  Query Match            100.0%;  Score 521;  DB 5;  Length 530;
  Best Local Similarity  100.0%;  Pred. No. 9e-45;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 7
US-09-782-051-2
; Sequence 2, Application US/09782051
; Patent No. US20020035078A1
; GENERAL INFORMATION:
;  APPLICANT: Hart, Derek N J
;  TITLE OF INVENTION: Enzyme having S-adenosyl-L-homocysteine hydrolase
;  TITLE OF INVENTION:  (AHCY) type activity
;  FILE REFERENCE: 24305 MRB
;  CURRENT APPLICATION NUMBER: US/09/782,051
;  CURRENT FILING DATE:  2001-02-14
;  PRIOR APPLICATION NUMBER: PCT/NZ97/00133
;  PRIOR FILING DATE: 1997-10-06
;  PRIOR APPLICATION NUMBER: NZ 299507
;  PRIOR FILING DATE: 1996-10-04
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 614
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-782-051-2

  Query Match            100.0%;  Score 521;  DB 3;  Length 614;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-44;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 144

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||||||||||||||||||||||||||||
Db        145 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 188


RESULT 8
US-10-450-763-42388
; Sequence 42388, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 42388
;   LENGTH: 455
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (356)..(388)
;   OTHER INFORMATION: S-adenosyl-L-homocysteine hydrolase proteins domain
;   OTHER INFORMATION: identified by eMATRIX, accession number BL00738F, p-value=7.261
;   OTHER INFORMATION: 29, raw score of 12.23
;   FEATURE:
;   NAME/KEY: DOMAIN
;   LOCATION: (186)..(447)
;   OTHER INFORMATION: S-adenosyl-L-homocysteine hydrolase domain identified by
;   OTHER INFORMATION: PFam, accession name AdoHcyase, E-value=6.2e-171, PFam score of
;   OTHER INFORMATION: 581.3
US-10-450-763-42388

  Query Match            55.4%;  Score 288.5;  DB 5;  Length 455;
  Best Local Similarity  54.0%;  Pred. No. 5.4e-21;
  Matches   67;  Conservative   13;  Mismatches   17;  Indels   27;  Gaps    3;

Qy          8 PLPGVGE-----------ELKQAKEIEDAEKY--------------SFMATVTKAPK--K 40
              |:|| |            :: ||  :: :: :              |   |||:||:  |
Db         61 PVPGPGSGPAAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 120

Qy         41 QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSK 100
              :|||||  ||| | ||| ||||||||||||||||||||||||||||| |||:||| |||
Db        121 KIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSK 180

Qy        101 GSSN 104
              |||:
Db        181 GSSD 184


RESULT 9
US-10-756-149-5569
; Sequence 5569, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS A
;  TITLE OF INVENTION:  METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
;  FILE REFERENCE: file
;  CURRENT APPLICATION NUMBER: US/10/756,149
;  CURRENT FILING DATE:  2004-01-12
;  NUMBER OF SEQ ID NOS: 5818
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5569
;   LENGTH: 619
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-756-149-5569
```

```
  Query Match             55.4%;  Score 288.5;  DB 5;  Length 619;
  Best Local Similarity   54.0%;  Pred. No. 8e-21;
  Matches   67;  Conservative   13;  Mismatches   17;  Indels   27;  Gaps    3;

Qy          8 PLPGVGE----------ELKQAKEIEDAEKY-------------SFMATVTKAPK--K 40
              |:|| |            :: ||  :: :: :              |   |||:||:  |
Db         70 PVPGPGSGPAAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 129

Qy         41 QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSK 100
              :|||||  ||| | ||| ||||||||||||||||||||||||||||| |||:||| |||
Db        130 KIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSK 189

Qy        101 GSSN 104
              |||:
Db        190 GSSD 193


RESULT 10
US-10-424-599-212758
; Sequence 212758, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 212758
;   LENGTH: 257
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(257)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_34146C.1.pep
US-10-424-599-212758

  Query Match             55.0%;  Score 286.5;  DB 4;  Length 257;
  Best Local Similarity   54.0%;  Pred. No. 4.2e-21;
  Matches   67;  Conservative   12;  Mismatches   18;  Indels   27;  Gaps    3;

Qy          8 PLPGVGE----------ELKQAKEIEDAEKY-------------SFMATVTKAPK--K 40
              | || |            :: ||  :: :: :              |   |||:||:  |
Db         63 PAPGPGSGPTAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 122

Qy         41 QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSK 100
              :|||||  ||| | ||| ||||||||||||||||||||||||||||| |||:||| |||
Db        123 KIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSK 182

Qy        101 GSSN 104
              |||:
Db        183 GSSD 186


RESULT 11
US-10-467-903-1
; Sequence 1, Application US/10467903
; Publication No. US20040265807A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE CORPORATION
;  APPLICANT:  SANJANWALA, Madhusudan M.
;  APPLICANT:  LU, Yan
;  APPLICANT:  LEE, Ernestine A.
;  APPLICANT:  HAFALIA, April J.A.
;  APPLICANT:  WARREN, Bridget A.
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  YUE, Henry
;  APPLICANT:  YAO, Monique G.
;  APPLICANT:  LEE, Sally
;  APPLICANT:  THORNTON, Michael B.
;  APPLICANT:  CHAWLA, Narinder K.
;  APPLICANT:  XU, Yuming
;  APPLICANT:  TRAN, Uyen K.
;  APPLICANT:  LAL, Preeti G.
;  APPLICANT:  LU, Dyung Aina M.
;  APPLICANT:  SWARNAKAR, Anita
;  APPLICANT:  RING, Huijun Z.
;  APPLICANT:  JONES, Karen A.
;  TITLE OF INVENTION: ENZYMES
;  FILE REFERENCE: PI-0360 USN
;  CURRENT APPLICATION NUMBER: US/10/467,903
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: PCT/US02/03814
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/268,113
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: US 60/269,215
;  PRIOR FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: US 60/272,271
;  PRIOR FILING DATE: 2001-02-27
;  PRIOR APPLICATION NUMBER: US 60/274,091
;  PRIOR FILING DATE: 2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/274,423
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: US 60/278,480
;  PRIOR FILING DATE: 2001-03-23
;  PRIOR APPLICATION NUMBER: US 60/278,479
;  PRIOR FILING DATE: 2001-03-23
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PERL Program
; SEQ ID NO 1
;   LENGTH: 564
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7484737CD1
US-10-467-903-1

  Query Match             52.4%;  Score 273;  DB 5;  Length 564;
  Best Local Similarity   73.1%;  Pred. No. 2.8e-19;
  Matches   57;  Conservative    4;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         27 KYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD 86
              |  ::      | ||||||  ||| | ||| ||||||||||||||||||||||||||||||
Db         61 KKKYIVNGNSGIKAQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSD 120

Qy         87 DEVSPREKQQTNSKGSSN 104
              || |||:||| ||||||:
Db        121 DETSPRDKQQKNSKGSSD 138


RESULT 12
US-11-097-143-2781
; Sequence 2781, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
```

```
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2781
;   LENGTH: 521
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-11-097-143-2781

  Query Match             24.3%;  Score 126.5;  DB 6;  Length 521;
  Best Local Similarity   64.0%;  Pred. No. 0.00028;
  Matches   32;  Conservative    5;  Mismatches   10;  Indels    3;  Gaps    2;

Qy         57 KTGRRSLSRSISQSSTDSYSSAASYTDSSD--DEVSPREKQQTNSKGSSN 104
              |    |  |||:| ||||:|| |||| ||:  |:| |||| | ||||||:
Db         46 KKTSRYRSRSLSASSTDSFSS-ASYTGSSEDGDDVPPREKVQKNSKGSSD 94


RESULT 13
US-11-004-418A-28
; Sequence 28, Application US/11004418A
; Publication No. US20050227322A1
; GENERAL INFORMATION:
;  APPLICANT: Lindquist et al.
;  TITLE OF INVENTION: RECOMBINANT PRION-LIKE GENES AND PROTEINS AND MATERIALS
;  TITLE OF INVENTION:  AND METHODS COMPRISING SAME
;  FILE REFERENCE: 30554/34978B
;  CURRENT APPLICATION NUMBER: US/11/004,418A
;  CURRENT FILING DATE:  2004-12-03
;  PRIOR APPLICATION NUMBER: US 06/138,833
;  PRIOR FILING DATE: 1999-06-09
;  NUMBER OF SEQ ID NOS: 65
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 414
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-11-004-418A-28

  Query Match             16.6%;  Score 86.5;  DB 6;  Length 414;
  Best Local Similarity   31.4%;  Pred. No. 2.6;
  Matches   32;  Conservative   16;  Mismatches   39;  Indels   15;  Gaps    3;

Qy         17 KQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYS 76
              ||||| | |:  |  :: : :       ::   |     : :   | | | | ||:|| |
Db         18 KQAKE-EKAKAVSSSSSESSSSSSSSSESESESESESSSSSSSSDSESSSSSSSDSES 76

Qy         77 ----------SAASYTDSSDDEVSPREKQQT---NSKGSSN 104
                      |:::| :||| ||    ||::|    || ||:
Db         77 EAETKKEESKDSSSSSSDSSSDEEEEEKEETKKEESKESSS 118


RESULT 14
US-10-424-599-169465
; Sequence 169465, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 169465
;   LENGTH: 321
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_124040C.1.pep
US-10-424-599-169465

  Query Match             16.2%;  Score 84.5;  DB 4;  Length 321;
  Best Local Similarity   29.3%;  Pred. No. 3.1;
  Matches   36;  Conservative   18;  Mismatches   28;  Indels   41;  Gaps    6;

Qy          8 PLPGVGEELK-----QAKEIE----------DAEKYSFMATVTKAPKKQIQFADDMQEFT 52
              |   | |: |      | |::|          :||| :  |||   | |
Db        124 PQKMVAEKQKNGATVQKKKVESSSDDSSSESEAEKPAVKATV---PSK------------ 168

Qy         53 KFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD-DEVSPRE----------KQQTNSKG 101
              | ||| |  |      :::: | || :|  |||| |||: ::          |::::|
Db        169 KTPTKNGNLSTPAKKGKAASSSSSSDSSEDDSSDEDEVATKKQTKEVKVQKGKEESSSDD 228

Qy        102 SSN 104
              ||:
Db        229 SSS 231


RESULT 15
US-10-052-648A-20
; Sequence 20, Application US/10052648A
; Publication No. US20040005558A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Colman, Steven
;  APPLICANT:  Edinger, Shlomit R.
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Mehraban, Fuad
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Rothenberg, Mark
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
;  TITLE OF INVENTION:  USING THE SAME
;  FILE REFERENCE: 21402-250 (CURA-550)
;  CURRENT APPLICATION NUMBER: US/10/052,648A
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/262,454
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/272,920
;  PRIOR FILING DATE: 2001-03-02
;  PRIOR APPLICATION NUMBER: 60/284,549
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/303,229
;  PRIOR FILING DATE: 2001-07-05
;  PRIOR APPLICATION NUMBER: 60/262,892
;  PRIOR FILING DATE: 2001-01-19
;  PRIOR APPLICATION NUMBER: 60/263,605
;  PRIOR FILING DATE: 2001-01-23
;  PRIOR APPLICATION NUMBER: 60/269,098
```

```
;  PRIOR FILING DATE: 2001-02-15
;  PRIOR APPLICATION NUMBER: 60/264,159
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/265,517
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: 60/271,855
;  PRIOR FILING DATE: 2001-02-27
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 97
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 2058
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-052-648A-20

  Query Match             16.0%;  Score 83.5;  DB 4;  Length 2058;
  Best Local Similarity   28.4%;  Pred. No. 41;
  Matches   31;  Conservative   19;  Mismatches   42;  Indels   17;  Gaps    3;

Qy          12 VGEELKQAKEIEDAEKYSFMATV-------TKAPKKQIQFADDMQEFTKFPTKTGR---- 60
               | |: |: |: |  |||   |:|       |  | : : | : :  :|   |: :
Db         851 VFEKHKEKKDKESTEKYKDRASVDSTQDKKNKQEKAEKKHAAEDKAKSKHKEKSDKEHSK 910

Qy          61 -RSLSRSISQSSTDSYSSAASYTDSSDDEVS-----PREKQQTNSKGSS 103
                |  |||      :|  |: |:||  ||: :      :||:  :|  ||
Db         911 ERKSSRSADAEYRESEVSSDSFTDREDDKSACLPEKLKEKRHRHSSSSS 959


Search completed: December 14, 2005, 08:31:27
Job time : 24.7855 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:21:03 ; Search time 1.64038 Seconds
(without alignments)
426.936 Million cell updates/sec

Title: US-10-683-610-1_COPY_1_104
Perfect score: 521
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....SDDEVSPREKQQTNSKGSSN 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 51441 seqs, 6734002 residues

Total number of hits satisfying chosen parameters: 51441

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	86.5	16.6	414	7	US-11-089-551A-28	Sequence 28, Appl
2	79	15.2	948	6	US-10-523-477-14	Sequence 14, Appl
3	70.5	13.5	99	6	US-10-995-561-1006	Sequence 1006, Ap
4	70.5	13.5	99	6	US-10-995-561-1007	Sequence 1007, Ap
5	69	13.2	440	7	US-11-082-389-106	Sequence 106, App
6	69	13.2	569	7	US-11-082-389-104	Sequence 104, App
7	68.5	13.1	1450	6	US-10-485-517-152	Sequence 152, App
8	67.5	13.0	761	6	US-10-485-517-252	Sequence 252, App
9	65	12.5	468	6	US-10-131-826A-90	Sequence 90, Appl
10	65	12.5	468	7	US-11-102-240-150	Sequence 150, App
11	65	12.5	486	7	US-11-108-172-685	Sequence 685, App
12	65	12.5	595	7	US-11-102-476-33	Sequence 33, Appl
13	65	12.5	1225	7	US-11-102-476-2	Sequence 2, Appli
14	64.5	12.4	896	6	US-10-467-657-7004	Sequence 7004, Ap
15	63.5	12.2	858	6	US-10-613-744-6	Sequence 6, Appli
16	63	12.1	273	7	US-11-196-475-98	Sequence 98, Appl
17	63	12.1	273	7	US-11-196-475-101	Sequence 101, App
18	63	12.1	274	7	US-11-196-475-9	Sequence 9, Appli
19	63	12.1	274	7	US-11-196-475-86	Sequence 86, Appl
20	63	12.1	274	7	US-11-196-475-138	Sequence 138, App
21	63	12.1	447	7	US-11-196-475-162	Sequence 162, App
22	63	12.1	454	7	US-11-196-475-160	Sequence 160, App
23	62.5	12.0	447	6	US-10-967-527A-14	Sequence 14, Appl
24	62.5	12.0	919	6	US-10-821-234-1144	Sequence 1144, Ap
25	62.5	12.0	964	7	US-11-089-551A-30	Sequence 30, Appl
26	62	11.9	140	6	US-10-793-626-2268	Sequence 2268, Ap
27	62	11.9	140	6	US-10-793-626-2438	Sequence 2438, Ap
28	62	11.9	710	7	US-11-089-551A-33	Sequence 33, Appl
29	61.5	11.8	273	7	US-11-196-475-11	Sequence 11, Appl
30	61.5	11.8	273	7	US-11-196-475-89	Sequence 89, Appl
31	61.5	11.8	273	7	US-11-196-475-140	Sequence 140, App
32	61.5	11.8	273	7	US-11-196-475-142	Sequence 142, App
33	61.5	11.8	296	7	US-11-078-280-2	Sequence 2, Appli
34	61.5	11.8	389	7	US-11-074-176-324	Sequence 324, App
35	61.5	11.8	406	7	US-11-074-176-92	Sequence 92, Appl
36	61.5	11.8	1448	6	US-10-485-517-212	Sequence 212, App
37	60.5	11.6	467	6	US-10-793-626-2652	Sequence 2652, Ap
38	60.5	11.6	1227	6	US-10-793-626-96	Sequence 96, Appl
39	60.5	11.6	1531	7	US-11-087-227-24	Sequence 24, Appl
40	60	11.5	655	6	US-10-793-626-1052	Sequence 1052, Ap
41	60	11.5	655	6	US-10-793-626-1400	Sequence 1400, Ap
42	60	11.5	743	6	US-10-485-517-351	Sequence 351, App
43	60	11.5	877	6	US-10-485-517-200	Sequence 200, App
44	60	11.5	1107	6	US-10-485-517-145	Sequence 145, App
45	59.5	11.4	448	6	US-10-967-527A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-11-089-551A-28
; Sequence 28, Application US/11089551A
; Publication No. US20050266242A1
; GENERAL INFORMATION:
; APPLICANT: Lindquist et al.
; TITLE OF INVENTION: ELECTRICAL CONDUCTORS AND DEVICES FROM PRION-LIKE PROTEINS
; FILE REFERENCE: 30554/40025A
; CURRENT APPLICATION NUMBER: US/11/089,551A
; CURRENT FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: US 60/559,286
; PRIOR FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-11-089-551A-28

Query Match 16.6%; Score 86.5; DB 7; Length 414;
Best Local Similarity 31.4%; Pred. No. 0.082;
Matches 32; Conservative 16; Mismatches 39; Indels 15; Gaps 3;

```
Qy         17 KQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYS 76
              ||||| | |:  |  :: : :        ::   |     : :   | | | | ||:|| |
Db         18 KQAKE-EKAKAVSSSSSESSSSSSSSSESESESESESESSSSSSSSDSESSSSSSSDSES 76

Qy         77 -----------SAASYTDSSDDEVSPREKQQT---NSKGSSN 104
                        |::| :||| ||    ||::|    || ||:
Db         77 EAETKKEESKDSSSSSSDSSSDEEEEEEKEETKKEESKESSS 118
```

RESULT 2
US-10-523-477-14
; Sequence 14, Application US/10523477
; Publication No. US20050266406A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MAXS AS MODIFIERS OF THE AXIN PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-051C-US
; CURRENT APPLICATION NUMBER: US/10/523,477
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US 60/401,534
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/411,153
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 14

```
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
;   LENGTH: 948
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-523-477-14

  Query Match             15.2%;  Score 79;  DB 6;  Length 948;
  Best Local Similarity   29.8%;  Pred. No. 1.2;
  Matches   31;  Conservative   17;  Mismatches   48;  Indels    8;  Gaps    4;

Qy          2 SMPDAMPLPGVGEEL-KQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ::|      |   ||  :: : ::: ||       :|  :    || :: |: |  |  |
Db        451 TIPTTQAPPLSREEFYREQRRLKEEEKKK-----SKLDEFTNDFAKELMEYKKI-QKERR 504

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              || ||| |  |  || | :||| |     | | :  : |   |:
Db        505 RSFSRSKSPYSGSSY-SRSSYTYSKSRSGSTRSRSYSRSFSRSH 547


RESULT 3
US-10-995-561-1006
; Sequence 1006, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1006
;   LENGTH: 99
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-995-561-1006

  Query Match             13.5%;  Score 70.5;  DB 6;  Length 99;
  Best Local Similarity   29.3%;  Pred. No. 0.59;
  Matches   24;  Conservative    9;  Mismatches   22;  Indels   27;  Gaps    3;

Qy         18 QAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSS------ 71
              :| | |||   ||                 || : |  ||| : :|| |: :|
Db         19 RASEAEDASLLSF----------------MQGYMKHATKTAKDALS-SVQESQVAQQAR 60

Qy         72 ---TDSYSSAASYTDSSDDEVS 90
                 || :||   |  :  |: |
Db         61 GWVTDGFSSLKDYWSTVKDKFS 82


RESULT 4
US-10-995-561-1007
; Sequence 1007, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1007
;   LENGTH: 99
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-995-561-1007

  Query Match             13.5%;  Score 70.5;  DB 6;  Length 99;
  Best Local Similarity   29.3%;  Pred. No. 0.59;
  Matches   24;  Conservative    9;  Mismatches   22;  Indels   27;  Gaps    3;

Qy         18 QAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSS------ 71
              :| | |||   ||                 || : |  ||| : :|| |: :|
Db         19 RASEAEDASLLSF----------------MQGYMKHATKTAKDALS-SVQESQVAQQAR 60

Qy         72 ---TDSYSSAASYTDSSDDEVS 90
                 || :||   |  :  |: |
Db         61 GWVTDGFSSLKDYWSTVKDKFS 82


RESULT 5
US-11-082-389-106
; Sequence 106, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
;  APPLICANT: Pompejus, Markus
;  APPLICANT:  Kroger, Burkhard
;  APPLICANT:  Schroder, Hartwig
;  APPLICANT:  Zelder, Oskar
;  APPLICANT:  Haberhauer, Gregor
;  TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
;  TITLE OF INVENTION:  INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
;  TITLE OF INVENTION:  TRANSPORT
;  FILE REFERENCE: BGI-131CPCN
;  CURRENT APPLICATION NUMBER: US/11/082,389
;  CURRENT FILING DATE:  2005-03-16
;  PRIOR APPLICATION NUMBER: US 09/603024
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/141031
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: US 60/143262
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: US 60/151281
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: DE 19930487.4
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: DE 19930489.0
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: DE 19931549.3
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931550.7
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19932134.5
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19941379.7
;  PRIOR FILING DATE: 1999-08-31
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 106
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-11-082-389-106

  Query Match             13.2%;  Score 69;  DB 7;  Length 440;
  Best Local Similarity   23.6%;  Pred. No. 5;
  Matches   17;  Conservative   20;  Mismatches   23;  Indels   12;  Gaps    3;

Qy          9 LPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTK----TGRRSLS 64
              :| : : |   ::: : :| |   ::  | | :: | |:       ||:     |::||:
Db        302 VPNIPDHL-HPRDLSEGQKLSLALSIQLAAKPRVVFFDE-------PTRGLDYDGKKSLA 353

Qy         65 RSISQSSTDSYS 76
              ||  | : | ::
Db        354 RSFQQLADDGHA 365


RESULT 6
US-11-082-389-104
; Sequence 104, Application US/11082389
```

```
; Publication No. US20050244935A1
; GENERAL INFORMATION:
;  APPLICANT: Pompejus, Markus
;  APPLICANT:  Kroger, Burkhard
;  APPLICANT:  Schroder, Hartwig
;  APPLICANT:  Zelder, Oskar
;  APPLICANT:  Haberhauer, Gregor
;  TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
;  TITLE OF INVENTION:  INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
;  TITLE OF INVENTION:  TRANSPORT
;  FILE REFERENCE: BGI-131CPCN
;  CURRENT APPLICATION NUMBER: US/11/082,389
;  CURRENT FILING DATE:  2005-03-16
;  PRIOR APPLICATION NUMBER: US 09/603024
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/141031
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: US 60/143262
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: US 60/151281
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: DE 19930487.4
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: DE 19930489.0
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: DE 19931549.3
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931550.7
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19932134.5
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19941379.7
;  PRIOR FILING DATE: 1999-08-31
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 104
;   LENGTH: 569
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-11-082-389-104

  Query Match             13.2%;  Score 69;  DB 7;  Length 569;
  Best Local Similarity   23.6%;  Pred. No. 6.8;
  Matches   17;  Conservative   20;  Mismatches   23;  Indels   12;  Gaps    3;

Qy          9 LPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTK----TGRRSLS 64
              :| : : |   ::: : :| |   ::  | | :: | |:       ||:     |::||:
Db        431 VPNIPDHL-HPRDLSEGQKLSLALSIQLAAKPRVVFFDE-------PTRGLDYDGKKSLA 482

Qy         65 RSISQSSTDSYS 76
              ||  | : | ::
Db        483 RSFQQLADDGHA 494


RESULT 7
US-10-485-517-152
; Sequence 152, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
;  APPLICANT: University of Sheffield
;  APPLICANT:  Biosynexus Incorporated
;  APPLICANT:  Foster, Simon
;  APPLICANT:  Mond, James
;  TITLE OF INVENTION: Antigenic Polypeptides
;  FILE REFERENCE: P100629WO
;  CURRENT APPLICATION NUMBER: US/10/485,517
;  CURRENT FILING DATE:  2004-02-02
;  PRIOR APPLICATION NUMBER: GB 0118825.9
;  PRIOR FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: GB 0200349.9
;  PRIOR FILING DATE: 2002-01-09
;  NUMBER OF SEQ ID NOS: 424
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 152
;   LENGTH: 1450
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-10-485-517-152

  Query Match             13.1%;  Score 68.5;  DB 6;  Length 1450;
  Best Local Similarity   26.5%;  Pred. No. 23;
  Matches   22;  Conservative   16;  Mismatches   34;  Indels   11;  Gaps    3;

Qy          5 DAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLS 64
              | ||:  : :     ::||| ||  |:|     :||::|     :   : |   |:::
Db         90 DTMPV--IQQVFIDIRDIEDVEKRLFLA------RKQLEFYSTQCDLELYFTSLSRKTI- 140

Qy         65 RSISQSSTDSYSSAASYTDSSDD 87
                : :    |      ||| |||
Db        141 --VYKGWLRSDQIKKLYTDLSDD 161


RESULT 8
US-10-485-517-252
; Sequence 252, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
;  APPLICANT: University of Sheffield
;  APPLICANT:  Biosynexus Incorporated
;  APPLICANT:  Foster, Simon
;  APPLICANT:  Mond, James
;  TITLE OF INVENTION: Antigenic Polypeptides
;  FILE REFERENCE: P100629WO
;  CURRENT APPLICATION NUMBER: US/10/485,517
;  CURRENT FILING DATE:  2004-02-02
;  PRIOR APPLICATION NUMBER: GB 0118825.9
;  PRIOR FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: GB 0200349.9
;  PRIOR FILING DATE: 2002-01-09
;  NUMBER OF SEQ ID NOS: 424
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 252
;   LENGTH: 761
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-10-485-517-252

  Query Match             13.0%;  Score 67.5;  DB 6;  Length 761;
  Best Local Similarity   25.0%;  Pred. No. 14;
  Matches   23;  Conservative   19;  Mismatches   37;  Indels   13;  Gaps    2;

Qy         24 DAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRR----------SLSRSISQSST 72
              ::|  |  :: :|:    :  :  | : |   | |              ||| |::||
Db        179 ESESDSTSSSESKSDSTSMSIS--MSQSTSGSTSTSTSTSLSDSTSTSLSLSASMNQSGV 236

Qy         73 DSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              || |:: | ::|:    |  : | |:|  | :
Db        237 DSNSASQSASNSTSTSTSESDSQSTSSYTSQS 268


RESULT 9
US-10-131-826A-90
; Sequence 90, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
```

```
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C128
;  CURRENT APPLICATION NUMBER: US/10/131,826A
;  CURRENT FILING DATE:  2002-04-24
;  PRIOR APPLICATION NUMBER: 60/049911
;  PRIOR FILING DATE: 1997-06-18
;  PRIOR APPLICATION NUMBER: 60/056974
;  PRIOR FILING DATE: 1997-08-26
;  PRIOR APPLICATION NUMBER: 60/059113
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059115
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059117
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059122
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059184
;  PRIOR FILING DATE: 1997-09-17
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059352
;  PRIOR FILING DATE: 1997-09-19
;  PRIOR APPLICATION NUMBER: 60/059588
;  PRIOR FILING DATE: 1997-09-19
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
;   LENGTH: 468
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-131-826A-90

  Query Match             12.5%;  Score 65;  DB 6;  Length 468;
  Best Local Similarity   23.6%;  Pred. No. 13;
  Matches   25;  Conservative   23;  Mismatches   30;  Indels   28;  Gaps    5;

Qy         17 KQAKEIEDAEKYSFMATV----------TKAPKKQIQFADDMQEFTKFPTKTGRRSLSR 65
              |:||| |     : | |:           | :|:: : : ::: |      ||
Db        291 KEAKEKETL--ITIMKTLIDFVKMMVKYGTISPEEGVSYLENLDEMIALQTK-------N 341

Qy         66 SISQSSTDSYS------SAASY--TDSSDDEVSPREKQQTNSKGSS 103
                : :::||: |       |  |:  |||: :| :  ||:  : | |:
Db        342 KLEKNATDNISKLFPAPSEKSHEETDSTKEEAAKMEKEYGSLKDST 387


RESULT 10
US-11-102-240-150
; Sequence 150, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
;  APPLICANT: Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Grimaldi,Christopher J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Wood,William I.
;  TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
;  TITLE OF INVENTION:  ESOPHAGEAL TUMOR
;  FILE REFERENCE: P3230R1C106C
;  CURRENT APPLICATION NUMBER: US/11/102,240
;  CURRENT FILING DATE:  2005-04-08
;  PRIOR APPLICATION NUMBER: 10/063662
;  PRIOR FILING DATE: 2002-05-07
;  PRIOR APPLICATION NUMBER: 10/006867
;  PRIOR FILING DATE: 2001-12-06
;  PRIOR APPLICATION NUMBER: PCT/US00/23328
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: 60/170262
;  PRIOR FILING DATE: 199-12-09
;  NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 150
;   LENGTH: 468
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-11-102-240-150

  Query Match             12.5%;  Score 65;  DB 7;  Length 468;
  Best Local Similarity   23.6%;  Pred. No. 13;
  Matches   25;  Conservative   23;  Mismatches   30;  Indels   28;  Gaps    5;

Qy         17 KQAKEIEDAEKYSFMATV----------TKAPKKQIQFADDMQEFTKFPTKTGRRSLSR 65
              |:|:|| |     : | |:           | :|:: : : ::: |      ||
Db        291 KEAKEKETL--ITIMKTLIDFVKMMVKYGTISPEEGVSYLENLDEMIALQTK-------N 341

Qy         66 SISQSSTDSYS------SAASY--TDSSDDEVSPREKQQTNSKGSS 103
                : :::||: |       |  |:  |||: :| :  ||:  : | |:
Db        342 KLEKNATDNISKLFPAPSEKSHEETDSTKEEAAKMEKEYGSLKDST 387


RESULT 11
US-11-108-172-685
; Sequence 685, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Clapper, Jonathan D.
;  APPLICANT:  Skeiky, Yasir A. W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick Thomas S.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
;  TITLE OF INVENTION:  OF COLON CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.471C15
;  CURRENT APPLICATION NUMBER: US/11/108,172
;  CURRENT FILING DATE:  2005-04-15
;  PRIOR APPLICATION NUMBER: US 10/025,380
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: US 09/922,217
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: US 09/833,263
;  PRIOR FILING DATE: 2001-04-10
;  PRIOR APPLICATION NUMBER: US 09/649,811
;  PRIOR FILING DATE: 2000-08-28
;  PRIOR APPLICATION NUMBER: US 09/609,448
;  PRIOR FILING DATE: 2000-06-29
;  PRIOR APPLICATION NUMBER: US 09/575,251
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: US 09/519,444
;  PRIOR FILING DATE: 2000-03-06
;  PRIOR APPLICATION NUMBER: US 09/504,629
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: US 09/480,321
;  PRIOR FILING DATE: 2000-01-10
;  PRIOR APPLICATION NUMBER: US 09/476,296
;  PRIOR FILING DATE: 1999-12-30
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1130
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 685
;   LENGTH: 486
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-108-172-685

  Query Match             12.5%;  Score 65;  DB 7;  Length 486;
  Best Local Similarity   26.2%;  Pred. No. 14;
  Matches   28;  Conservative   11;  Mismatches   50;  Indels   18;  Gaps    6;

Qy          5 DAMPLPGV---GEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTK--TG 59
              |   | |:   |::   :|           ::    |    | |    :  |  :  ||
Db        322 DGFTLTGINQTGDQFLPSKP----------SSTLNVPSNPAQSAQSQSSYNPFEDEDDTG 371

Qy         60 RR-SLSRSISQSSTDSYSSAASY-TDSSDDEVSPREKQQTNSKGSSN 104
                |   |   :  ||    || || |||| |      |:: | ||
Db        372 STVSEKEDIKAKNVSSYEKTQSYPTDWSDDE-SNNPFSSTDANGDSN 417


RESULT 12
US-11-102-476-33
; Sequence 33, Application US/11102476
; Publication No. US20050271680A1
; GENERAL INFORMATION:
;  APPLICANT: Chang, Yung-Fu
;  APPLICANT:  Palaniappan, Raghavan U.M.
;  APPLICANT:  Cornell Research Foundation, Inc.
;  TITLE OF INVENTION: Novel Immunogenic Proteins of Leptospira
;  FILE REFERENCE: 1153.080US1
;  CURRENT APPLICATION NUMBER: US/11/102,476
;  CURRENT FILING DATE:  2005-04-08
;  PRIOR APPLICATION NUMBER: PCT/US2003/32385
;  PRIOR FILING DATE: 2003-10-10
;  PRIOR APPLICATION NUMBER: US 60/417721
;  PRIOR FILING DATE: 2002-10-10
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
;   LENGTH: 595
;   TYPE: PRT
;   ORGANISM: Leptospira interrogans
US-11-102-476-33

  Query Match             12.5%;  Score 65;  DB 7;  Length 595;
  Best Local Similarity   23.8%;  Pred. No. 18;
  Matches   24;  Conservative   19;  Mismatches   30;  Indels   28;  Gaps    4;

Qy         11 GVGEELKQAK--------EIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRS 62
              |:|  |||             |||:  :    |||:|    |: :          ||:
Db         11 GLGSTLKQGTVKVTASMGGIEDSVDF----TVTQATLTSIEVS---------PTRA---- 53

Qy         63 LSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSS 103
                ||::  |  :::   :|| |   :: :   :::||  |
Db         54 ---SIAKGMTQKFTATGIFTDHSKKNITEQVTWKSSSKALS 91


RESULT 13
US-11-102-476-2
; Sequence 2, Application US/11102476
; Publication No. US20050271680A1
; GENERAL INFORMATION:
;  APPLICANT: Chang, Yung-Fu
;  APPLICANT:  Palaniappan, Raghavan U.M.
;  APPLICANT:  Cornell Research Foundation, Inc.
;  TITLE OF INVENTION: Novel Immunogenic Proteins of Leptospira
;  FILE REFERENCE: 1153.080US1
;  CURRENT APPLICATION NUMBER: US/11/102,476
;  CURRENT FILING DATE:  2005-04-08
;  PRIOR APPLICATION NUMBER: PCT/US2003/32385
;  PRIOR FILING DATE: 2003-10-10
;  PRIOR APPLICATION NUMBER: US 60/417721
;  PRIOR FILING DATE: 2002-10-10
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1225
;   TYPE: PRT
;   ORGANISM: Leptospira interrogans
US-11-102-476-2

  Query Match             12.5%;  Score 65;  DB 7;  Length 1225;
  Best Local Similarity   23.8%;  Pred. No. 43;
  Matches   24;  Conservative   19;  Mismatches   30;  Indels   28;  Gaps    4;

Qy         11 GVGEELKQAK--------EIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRS 62
              |:|  |||             |||:  :    |||:|    |: :          ||:
Db        641 GLGSTLKQGTVKVTASMGGIEDSVDF----TVTQATLTSIEVS---------PTRA---- 683

Qy         63 LSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSS 103
                ||::  |  :::   :|| |   :: :   :::||  |
Db        684 ---SIAKGMTQKFTATGIFTDHSKKNITEQVTWKSSSKALS 721


RESULT 14
US-10-467-657-7004
; Sequence 7004, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
;  APPLICANT: CHIRON SpA
;  APPLICANT: FONTANA Maria Rita
;  APPLICANT: PIZZA Mariagrazia
;  APPLICANT: MASIGNANI Vega
;  APPLICANT: MONACI Elisabetta
;  TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/467,657
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: GB-0103424.8
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 9218
;  SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7004
;   LENGTH: 896
;   TYPE: PRT
;   ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7004

  Query Match             12.4%;  Score 64.5;  DB 6;  Length 896;
  Best Local Similarity   25.0%;  Pred. No. 33;
  Matches   29;  Conservative   23;  Mismatches   33;  Indels   31;  Gaps    6;

Qy          6 AMPLPGVGEELKQAKEI-----EDAE----KYSFMATVTKAPKKQIQFADDMQEFTKFPT 56
              |: | |  | ||:  |:     |: |     | |:  |    : ::|: || |  |   :
Db        314 ALQLKGTQETLKKIAEMGIMPSENVETLLAAYRFLRDV----EHRLQYWDDRQTQTLPIS 369

Qy         57 KTGRRSLSRSISQSSTDSYSSAASYTDSSD----------DEV--SPREKQQTNSK 100
                 |: |: |:       : | |:::|  :          :|:   | |: |:||:
Db        370 PEQRQLLAESM------GFDSYAAFSDGLNVHRNKVNQLFNEILSEPEEQAQSNSE 419


RESULT 15
US-10-613-744-6
; Sequence 6, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
;  APPLICANT: MacKinnon, Roderick
;  APPLICANT:  The Rockefeller University
;  TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
;  TITLE OF INVENTION:  Cation Channel Proteins, Mutant Prokaryotic Cation
;  TITLE OF INVENTION:  Channel Proteins, and Uses Thereof
;  FILE REFERENCE: 018512-002901US
;  CURRENT APPLICATION NUMBER: US/10/613,744
;  CURRENT FILING DATE:  2003-07-03
;  PRIOR APPLICATION NUMBER: US/09/275,252
;  PRIOR FILING DATE: 1999-03-24
;  PRIOR APPLICATION NUMBER: US 09/045,529
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: US 09/054,347
```

```
;  PRIOR FILING DATE: 1998-04-02
;  PRIOR APPLICATION NUMBER: WO PCT/US99/06307
;  PRIOR FILING DATE: 1999-03-22
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 858
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-613-744-6

  Query Match             12.2%;  Score 63.5;  DB 6;  Length 858;
  Best Local Similarity   24.4%;  Pred. No. 39;
  Matches   32;  Conservative   15;  Mismatches   51;  Indels   33;  Gaps    4;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQ---------------IQFA 45
              |:   : |:    |   |:|  |: |  |  : |   | :                |  |
Db        536 MAKTQSQPILNTKESAAQSKPKEELEMESIPSPVAPLPTRTEGVIDMRSMSSIDSFISCA 595

Qy         46 DDMQEFTKF--------PTKTG---------RRSLSRSISQSSTDSYSSAASYTDSSDDE 88
               |  | |:|        |:|||         | :|  |  :    : |  ||   |   |
Db        596 TDFPEATRFSHSPLTSLPSKTGGSTAPEVGWRGALGASGGRFVEANPSPDASQHSSFFIE 655

Qy         89 VSPREKQQTNS 99
               ||:   :||:
Db        656 -SPKSSMKTNN 665


Search completed: December 14, 2005, 08:31:43
Job time : 2.64038 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:11:12 ; Search time 6.39748 Seconds
(without alignments)
1564.139 Million cell updates/sec

Title: US-10-683-610-1_COPY_1_104
Perfect score: 521
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....SDDEVSPREKQQTNSKGSSN 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	322	61.8	597	2	T08681	adenosylhomocystei
2	87	16.7	971	2	T19431	hypothetical prote
3	86.5	16.6	414	2	A39205	nuclear localizati
4	85.5	16.4	2717	2	A34203	DNA-binding protei
5	85	16.3	1788	2	T31095	vitellogenin precu
6	83.5	16.0	2038	2	A43742	female sterile hom
7	83	15.9	628	2	T39494	kinesin-like prote
8	82	15.7	302	2	F70220	conserved hypothet
9	81.5	15.6	542	2	S64030	probable membrane
10	81.5	15.6	788	2	T29768	hypothetical prote
11	81.5	15.6	915	2	T16623	hypothetical prote
12	81	15.5	368	2	G88636	protein W09G12.7 [
13	80	15.4	697	2	T16908	hypothetical prote
14	79.5	15.3	1791	2	T02345	hypothetical prote
15	79	15.2	948	2	A57640	retinoblastoma bin
16	78.5	15.1	471	2	T33997	hypothetical prote
17	78	15.0	416	2	G83656	hypothetical prote
18	78	15.0	1067	2	T06312	hypothetical prote
19	77.5	14.9	651	2	T46050	hypothetical prote
20	77	14.8	406	2	S38170	SRP40 protein - ye
21	77	14.8	557	2	A96527	probable nuM1 prot
22	77	14.8	640	2	S48423	hypothetical prote
23	76.5	14.7	577	2	S30237	transcription init
24	76	14.6	233	2	T00966	hypothetical prote
25	76	14.6	734	2	G86171	hypothetical prote
26	76	14.6	754	1	JC5314	CDC28/cdc2-like ki
27	75.5	14.5	635	2	T09648	nucleolin homolog
28	75.5	14.5	860	2	JC4566	chitinase (EC 3.2.
29	75.5	14.5	1359	2	T34036	hypothetical prote
30	75.5	14.5	1507	2	B47328	natural killer cel
31	75	14.4	354	2	T06033	hypothetical prote
32	75	14.4	389	2	S08031	nucleocapsid prote
33	75	14.4	500	2	S55785	nucleolar protein
34	75	14.4	527	1	S01302	adenosylhomocystei
35	75	14.4	990	2	I51618	nucleolar phosphop
36	74.5	14.3	200	2	T25406	hypothetical prote
37	74.5	14.3	772	2	T06154	hypothetical prote
38	74.5	14.3	2688	2	I49477	alpha-A-crystallin
39	74	14.2	245	2	S15076	protein kinase reg
40	74	14.2	245	2	S59927	14-3-3 protein, th
41	74	14.2	309	2	H70381	ferrochelatase (EC
42	74	14.2	389	1	VHIH2E	nucleocapsid prote
43	74	14.2	516	2	T40181	conserved hypothet
44	74	14.2	5327	2	T13564	microtubule-associ
45	73.5	14.1	380	2	G69604	spore coat protein

ALIGNMENTS

RESULT 1
T08681
adenosylhomocysteinase (EC 3.3.1.1) DKFZp564A1523 - human (fragment)
N;Alternate names: protein DKFZp564A1523.1
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08681
R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16469 .
A;Accession: T08681
A;Molecule type: mRNA
A;Residues: 1-597 <WAM>
A;Cross-references: UNIPROT:Q9UG84; UNIPARC:UPI000006FAA4; EMBL:AL049954
A;Experimental source: fetal brain; clone DKFZp564A1523
C;Genetics:
A;Note: DKFZp564A1523.1
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, human SAHH type
C;Keywords: NAD; thioether hydrolase
F;379-408/Region: beta-alpha-beta NAD nucleotide-binding fold
F;244,278/Active site: Cys #status predicted

Query Match 61.8%; Score 322; DB 2; Length 597;
Best Local Similarity 54.7%; Pred. No. 5.2e-22;
Matches 76; Conservative 8; Mismatches 19; Indels 36; Gaps 4;

```
Qy          1 MSMPDAMPLP-------GVGE---ELKQAKEIEDAEKYSFMATVTKAPKK---------- 40
              :| | |:|||       | |    |:| :  | ::    :  :  |
Db         34 ISQP-ALPLPQETQAPSGESENLRNLRQGRRAESSQNLLSLRRLVSAQGNGRSDWRWLWI 92

Qy         41 --------------QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSS 85
                            ||||||||||||||||||||||||||||||||||||||||||||
Db         93 QLWPEIDSLEISCFLQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSS 152

Qy         86 DDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||
Db        153 DDEVSPREKQQTNSKGSSN 171
```

RESULT 2
T19431
hypothetical protein C25A1.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19431
R;Mortimore, B.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19124
A;Accession: T19431
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-971 <WIL>

```
A;Cross-references: UNIPROT:Q9XVS4; UNIPARC:UPI000007C0A0; EMBL:Z81038; PIDN:CAB02755.1;
A;Experimental source: clone C25A1
C;Genetics:
A;Gene: CESP:C25A1.10
A;Map position: 1
A;Introns: 38/3; 92/3; 201/3; 919/3
C;Superfamily: nucleolus-cytoplasm shuttle phosphoprotein

  Query Match             16.7%;  Score 87;  DB 2;  Length 971;
  Best Local Similarity   27.4%;  Pred. No. 4.2;
  Matches   37;  Conservative   17;  Mismatches   43;  Indels   38;  Gaps    6;

Qy          4 PDAMPLPG------VGEELKQAKEIEDAEK------YSFMATVTKAPKK---QIQFADDM 48
              | | | |       | :: : : :  | ||       |  ||   | ||      |:||
Db        506 PAAKPTPANATPKPVAKKAESSSDSSDDEKKPVAKPTSAKATPKPAAKKADLSSDFSDDE 565

Qy         49 QEFTKFPTKTGRRSLSRSISQSSTDS--------------------YSSAASYTDSSDD 87
                | | |   :  |:  ::||:||                    :|:| :|||||
Db        566 APAKKTPAKAAPKPASKK-AESSSDSSDDEKPAAKSTPAKTTPKPTAKKAASSSSDSSDD 624

Qy         88 EVSPREKQQTNSKGS 102
              |  |  |  |::| :
Db        625 EKKPVAK-PTSAKAT 638


RESULT 3
A39205
nuclear localization sequence-binding protein NSR1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: nucleolar protein NSR1; p67 protein; protein G7001; protein YGR159c;
C;Species: Saccharomyces cerevisiae
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 31-Dec-2004
C;Accession: A39205; PC2395; S60450; S64468; S64470
R;Lee, W.C.; Xue, Z.; Melese, T.
J. Cell Biol. 113, 1-12, 1991
A;Title: The NSR1 gene encodes a protein that specifically binds nuclear localization se
A;Reference number: A39205; MUID:91177946; PMID:1706724
A;Accession: A39205
A;Molecule type: DNA
A;Residues: 1-414 <LEE>
A;Cross-references: UNIPROT:P27476; UNIPARC:UPI0000035CF2; GB:X57185; NID:g4057; PIDN:CA
R;Gamberi, C.; Contreas, G.; Romanelli, M.G.; Morandi, C.
Gene 148, 59-66, 1994
A;Title: Analysis of the yeast Nsr1 gene and protein domain comparison between Nsr1 and
A;Reference number: PC2395; MUID:95011659; PMID:7926838
A;Accession: PC2395
A;Molecule type: mRNA
A;Residues: 155-414 <GAM>
A;Cross-references: UNIPARC:UPI000017A377; EMBL:X57185
A;Note: The authors translated the codon TTC for residue 188 as Arg, GGT for residue 254
R;Skala, J.; Nawrocki, A.; Goffeau, A.
Yeast 11, 1421-1427, 1995
A;Title: The sequence of a 27 kb segment on the right arm of chromosome VII from Sacchar
.
A;Reference number: S60435; MUID:96158062; PMID:8585325
A;Accession: S60450
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-414 <SKA>
A;Cross-references: UNIPARC:UPI0000035CF2; EMBL:X85807; NID:g1045249; PIDN:CAA59817.1; P
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995
R;Van Dyck, L.; Skala, J.; de Wergifosse, P.; Purnelle, B.; Talla, E.; Nawrocki, A.; Del
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64428
A;Accession: S64468
A;Molecule type: DNA
A;Residues: 1-414 <VAN>
A;Cross-references: UNIPARC:UPI0000035CF2; EMBL:Z72944; NID:g1323270; PIDN:CAA97173.1; P
A;Experimental source: strain S288C
R;Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64071
A;Accession: S64470
A;Molecule type: DNA
A;Residues: 1-188 <RIE>
A;Cross-references: UNIPARC:UPI0000168F38; EMBL:Z72944; MIPS:YGR159c
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:NSR1
A;Cross-references: SGD:S0003391; MIPS:YGR159c
A;Map position: 7R
C;Keywords: DNA binding; nucleus; RNA binding
F;169-236/Domain: ribonucleoprotein repeat homology <RRM1>
F;268-336/Domain: ribonucleoprotein repeat homology <RRM2>

  Query Match             16.6%;  Score 86.5;  DB 2;  Length 414;
  Best Local Similarity   31.4%;  Pred. No. 1.7;
  Matches   32;  Conservative   16;  Mismatches   39;  Indels   15;  Gaps    3;

Qy         17 KQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYS 76
              ||||| | |:  |  :: : :       ::   |      : :   | | | | ||:|| |
Db         18 KQAKE-EKAKAVSSSSSESSSSSSSSSESESESESESSSSSSSSDSESSSSSSSDSES 76

Qy         77 ----------SAASYTDSSDDEVSPREKQQT---NSKGSSN 104
                      |::| :||| ||    ||::|    || ||:
Db         77 EAETKKEESKDSSSSSSDSSSDEEEEEEKEETKKEESKESSS 118


RESULT 4
A34203
DNA-binding protein PRDII-BF1 - human
N;Alternate names: major histocompatibility complex enhancer-binding protein 1
C;Species: Homo sapiens (man)
C;Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004
C;Accession: A34203; A34779
R;Fan, C.M.; Maniatis, T.
Genes Dev. 4, 29-42, 1990
A;Title: A DNA-binding protein containing two widely separated zinc finger motifs that
A;Reference number: A34203; MUID:90169514; PMID:2106471
A;Accession: A34203
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2717 <FAN>
A;Cross-references: UNIPROT:P15822; UNIPARC:UPI000013C3A1; EMBL:X51435; NID:g38017; PID
R;Baldwin Jr., A.S.; LeClair, K.P.; Singh, H.; Sharp, P.A.
Mol. Cell. Biol. 10, 1406-1414, 1990
A;Title: A large protein containing zinc finger domains binds to related sequence eleme
A;Reference number: A34779; MUID:90205817; PMID:2108316
A;Accession: A34779
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 801-1072,'N',1074-1168,'K',1170-1225,'V',1227-1434,'N',1436-1607,'I',1609-1
A;Cross-references: UNIPARC:UPI0000178663; GB:M32019
C;Superfamily: HIV-EP2 enhancer-binding protein
C;Keywords: DNA binding; transcription regulation; zinc finger

  Query Match             16.4%;  Score 85.5;  DB 2;  Length 2717;
  Best Local Similarity   32.6%;  Pred. No. 19;
  Matches   28;  Conservative   15;  Mismatches   32;  Indels   11;  Gaps    3;

Qy         18 QAKEIEDAEKYSF---MATVTKAP-------KKQIQFADDMQEFTKFP-TKTGRRSLSRS 66
              |  : | :||:|:     |::| |       ||::: |:     |:     |  |||||
Db       1249 QCHDQEKSEKFSWPQRSETLSKLPTEKLPPKKKRLRLAEIEHSSTESSFDSTLSRSLSRE 1308

Qy         67 ISQSSTDSYSSAASYTDSSDDEVSPR 92
              | | | |:|::    | |  | ||:
Db       1309 SSLSHTSSFSASLDIEDVSKTEASPK 1334


RESULT 5
T31095
vitellogenin precursor - Oreochromis aureus
C;Species: Oreochromis aureus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: T31095
R;Lim, E.H.; Lam, T.J.; Ding, J.L.
submitted to the EMBL Data Library, August 1997
```

```
A;Description: Cloning of full-length Oreochromis aureus vitellogenin cDNA and its deduc
A;Reference number: Z20978
A;Accession: T31095
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1788 <LIM>
A;Cross-references: UNIPROT:Q9YGK0; UNIPARC:UPI00000FE0A7; EMBL:AF017250; NID:g4102880;
C;Superfamily: vitellogenin

  Query Match             16.3%;  Score 85;  DB 2;  Length 1788;
  Best Local Similarity   44.4%;  Pred. No. 13;
  Matches   20;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         60 RRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||| ||| | ||: | ||::| : ||    : :  : :||| ||:
Db       1220 RRSRSRSSSSSSSSSSSSSSSSSSSSSSKNSKRSKSSNSKSSSS 1264


RESULT 6
A43742
female sterile homeotic protein, 205K - fruit fly (Drosophila melanogaster)
N;Alternate names: membrane protein fsh, 205K
N;Contains: female sterile homeotic protein, 110K
C;Species: Drosophila melanogaster
C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 31-Dec-2004
C;Accession: A43742; B43742
R;Haynes, S.R.; Mozer, B.A.; Bhatia-Dey, N.; Dawid, I.B.
Dev. Biol. 134, 246-257, 1989
A;Title: The Drosophila fsh locus, a maternal effect homeotic gene, encodes apparent mem
A;Reference number: A43742; MUID:89276730; PMID:2567251
A;Accession: A43742
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2038 <HAY>
A;Cross-references: UNIPROT:P13709; UNIPARC:UPI000012AC6C; EMBL:M23221; NID:g157452; PID
A;Accession: B43742
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1106 <HA2>
A;Cross-references: UNIPARC:UPI000017A10F; EMBL:M23222
C;Genetics:
A;Gene: fsh
A;Cross-references: FlyBase:FBgn0004656
C;Keywords: alternative splicing; transmembrane protein
F;1-2038/Product: female sterile homeotic protein, 205K #status predicted <MA2>
F;1-1106/Product: female sterile homeotic protein, 110K #status predicted <MAT>
F;59-116/Domain: bromodomain homology <BRO1>
F;503-560/Domain: bromodomain homology <BRO2>

  Query Match             16.0%;  Score 83.5;  DB 2;  Length 2038;
  Best Local Similarity   28.1%;  Pred. No. 21;
  Matches   39;  Conservative   16;  Mismatches   45;  Indels   39;  Gaps    6;

Qy          2 SMPDAMPLPGVGEELKQA--KEIEDAEKYSFMATVTKAP---KKQIQFADDMQEFTK-FP 55
              | || : :    | || :  :|:|          | |    |  | |:  || |
Db        989 SNPDEIEID--FETLKPSTLRELESYVASCLRKKTHKKPSGKSKDEQMAEKKQELEKRLQ 1046

Qy         56 TKTGRRSLSR----------------------SISQSSTDSYSSAASYTDSSD------- 86
                ||:   |:                     | | ||:|| ||::| : |||
Db       1047 DVTGQLGASKKTAKKDESASSKVEAVQPANPVSSSSSSSDSSSSSSSDSSSSDSSDSEAG 1106

Qy         87 --DEVSPREKQQTNSKGSS 103
                ||  ||:|:  :| ||:
Db       1107 DGDERPPRKKKSRDSNGSN 1125


RESULT 7
T39494
kinesin-like protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39494
R;Beck, A.; Reinhardt, R.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21859
A;Accession: T39494
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-628 <BEC>
A;Cross-references: UNIPROT:O94732; UNIPARC:UPI000006C858; EMBL:AL034433; PIDN:CAA2235
A;Experimental source: strain 972h-; cosmid c1604
C;Genetics:
A;Gene: SPDB:SPBC1604.20c
A;Map position: 1

  Query Match             15.9%;  Score 83;  DB 2;  Length 628;
  Best Local Similarity   28.6%;  Pred. No. 5.9;
  Matches   24;  Conservative   21;  Mismatches   23;  Indels   16;  Gaps    3;

Qy         14 EELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTD 73
              :||:  ::|  | :|:         :||   ||  |   ||| |  :::::|  |:||:|
Db        548 DELQDKEQIIYALRYA---------QKQRDIADFNQSLAKFPHKILKKNVTRG-SRSSSD 597

Qy         74 SYSSAASYTDSSDDEVSPREKQQT 97
              :|      : :  |: | ::||:
Db        598 QFS------NETKTEILPDDQQQS 615


RESULT 8
F70220
conserved hypothetical protein BBC05 - Lyme disease spirochete plasmid C/cp9
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C;Accession: F70220
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whit
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: F70220
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-302 <KLE>
A;Cross-references: UNIPROT:O50970; UNIPARC:UPI000005694A; GB:AE000791; NID:g3253099; P
A;Experimental source: strain B31
C;Genetics:
A;Genome: plasmid

  Query Match             15.7%;  Score 82;  DB 2;  Length 302;
  Best Local Similarity   28.0%;  Pred. No. 3.1;
  Matches   28;  Conservative   13;  Mismatches   49;  Indels   10;  Gaps    3;

Qy          7 MPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKT--GRRSLS 64
              | |||: : ||           ||     : :  || |    :   :| :||| |  | : :
Db         28 MNLPGIDKNLKGY-----GYKYQNFNEIAREIKKVIDKHNLCLDFKQFPTFTVVGEQQVL 82

Qy         65 RSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
                :    |  ||:   | || |  :     :  |  || |
Db         83 HVV---RTTFYSTNTGYKDSFDTPILTENLKWNNENGSKN 119


RESULT 9
S64030
probable membrane protein YGL028c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G3661
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 09-Jul-2004
C;Accession: S64030
R;Hebling, U.; Hofmann, B.; Delius, H.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64003
A;Accession: S64030
A;Molecule type: DNA
A;Residues: 1-542 <HEB>
```

```
A;Cross-references: UNIPROT:P53189; UNIPARC:UPI00001356BA; EMBL:Z72550; NID:g1322499; PI
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:SCW11; MIPS:YGL028c
A;Cross-references: SGD:S0002996
A;Map position: 7L
C;Keywords: transmembrane protein
F;8-24/Domain: transmembrane #status predicted <TMM>

  Query Match             15.6%;  Score 81.5;  DB 2;  Length 542;
  Best Local Similarity   26.7%;  Pred. No. 6.9;
  Matches   27;  Conservative   20;  Mismatches   45;  Indels    9;  Gaps    2;

Qy          11 GVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTK-------FPTKTGRRSL 63
               ||||   :      :: :|: |  |    :|   :   || |:         | :   |
Db         174 GVGEANVEGN--TQSQHHSYQAAATSTLNQQTSTSIASQESTESTNTPTSSSTSSSTSSS 231

Qy          64 SRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
               : | : ||| | :|::: : :|    | :|   | |:|||:
Db         232 TSSSTSSSTSSSTSSSTSSSTSSSTSSTQETAATTSEGSSS 272


RESULT 10
T29768
hypothetical protein ZC581.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Dec-2004
C;Accession: T29768
R;Waterston, B.; Gattung, S.; Le, T.T.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of C. elegans cosmid ZC581.
A;Reference number: Z20682
A;Accession: T29768
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-788 <WAT>
A;Cross-references: UNIPROT:O01772; UNIPARC:UPI0000179454; EMBL:AF003134; PIDN:AAB54141.
A;Experimental source: strain Bristol N2; clone ZC581
C;Genetics:
A;Gene: CESP:ZC581.3
A;Map position: 1
A;Introns: 29/3; 66/3; 90/2; 109/2; 144/1; 223/2; 282/3; 351/3; 426/1; 590/3; 699/1

  Query Match             15.6%;  Score 81.5;  DB 2;  Length 788;
  Best Local Similarity   25.7%;  Pred. No. 11;
  Matches   28;  Conservative   19;  Mismatches   39;  Indels   23;  Gaps    4;

Qy          14 EELKQAKE----IEDAEKYSFMATVTKAPKKQIQFADDMQE---FTK---FPTKTGRRSL 63
               || || :|    |:| | |     : |  ::     :||::   :||    ||:    ||
Db         263 EERKQEEERRRRIDDDEDYELSLRILKVQEELTTIVNDMEDDSSYTKTSTMPTEETTNSL 322

Qy          64 SRSISQSSTDSYSSAAS-------------YTDSSDDEVSPREKQQTNS 99
               : |:| |   :|: |             :|   || ::  |:  | :
Db         323 ESTSSESFTAESTSSISPILGTLESIKQEDFTTLLDDLITSAERDPTET 371


RESULT 11
T16623
hypothetical protein M02D8.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16623
R;Miller, N.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid M02D8.
A;Reference number: Z18549
A;Accession: T16623
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-915 <MIL>
A;Cross-references: UNIPROT:Q21461; UNIPARC:UPI0000081E0A; EMBL:U41034; NID:g1086742; PI
C;Genetics:
A;Gene: CESP:M02D8.3
A;Introns: 29/2; 85/2; 102/2; 142/2; 189/2; 214/2; 230/2; 243/2; 286/2; 356/2; 610/1; 6

  Query Match             15.6%;  Score 81.5;  DB 2;  Length 915;
  Best Local Similarity   24.6%;  Pred. No. 13;
  Matches   28;  Conservative   15;  Mismatches   50;  Indels   21;  Gaps    4;

Qy           3 MPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQ------IQFADDMQEFTKFPT 56
               || |:|:|   :  |  ||       |    : ::   |       :    ||   :  |
Db         298 MPSALPVP-TAQSSKSPKESSKEPPKSLCPAINRSDLPQTPQLAKVDENTDMNSSSDLPR 356

Qy          57 KTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS------PREKQQTNSKGSSN 104
               || |||  |   :   ||       |  :|:          :||::|  |  :|
Db         357 KTPRRSCVRGTYKDLVDS--------DGEEDKEKKTKMEMEKEKEKTRKKEENN 402


RESULT 12
G88636
protein W09G12.7 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: G88636
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_el
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A;Accession: G88636
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-368 <STO>
A;Cross-references: UNIPROT:O45198; UNIPARC:UPI000008159B; GB:chr_IV; PIDN:AAC04447.1;
C;Genetics:
A;Gene: W09G12.7
A;Map position: 4

  Query Match             15.5%;  Score 81;  DB 2;  Length 368;
  Best Local Similarity   24.1%;  Pred. No. 4.9;
  Matches   21;  Conservative   17;  Mismatches   31;  Indels   18;  Gaps    1;

Qy          36 KAPKKQIQFADDMQEFTKFPT-----------------KTGRRSLSRSISQSSTDSYSS 77
               |  ||  :  ||::   : |                 | |::|     | |:|  |
Db          25 KKDKKTFEEEDDVENVKRKPAESSDDEVVVVVVVKKDQKGKRGKKSKKPKKKQESSDDSSD 84

Qy          78 AASYTDSSDDEVSPREKQQTNSKGSSN 104
                  :  ::||||:   ::|:: : | ||:
Db          85 NSDSSESSDDDKKSKKKKKKSKKDSSD 111


RESULT 13
T16908
hypothetical protein T20B12.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-Oct-2004
C;Accession: T16908
R;Du, Z.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of C. elegans cosmid T20B12.
A;Reference number: S46772
A;Accession: T16908
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-697 <DUZ>
A;Cross-references: UNIPROT:P41848; UNIPARC:UPI000013600C; EMBL:U10401; NID:g500713; PI
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:T20B12.8
A;Introns: 79/3; 292/2; 601/3

  Query Match             15.4%;  Score 80;  DB 2;  Length 697;
  Best Local Similarity   26.3%;  Pred. No. 13;
  Matches   25;  Conservative   23;  Mismatches   27;  Indels   20;  Gaps    4;
```

```
Qy          23 EDAEKYSFMATVTKAPKKQIQFADDMQEFTKF--PTKTGRRSLSRSISQSSTDSYSSAAS 80
               :| :|:       || : : ::  :|:|: |   |: : : | |::  :||  | | | |
Db         598 DDKKKWE-----EKAEEDKSRYEKEMKEYRKNGPPSSSSKPSSSKTSKKSSGPSSSKAIS 652

Qy          81 Y-----TDSSDDE--------VSPREKQQTNSKGS 102
                  :| ||||         :|:|: : :: ||
Db         653 KEYISDSDDSDDEEPKKKEKKAAPKEESEESNNGS 687


RESULT 14
T02345
hypothetical protein KIAA0324 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C;Accession: T02345
R;Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
re, J.; White, S.; Ueng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
submitted to the EMBL Data Library, March 1998
A;Description: Sequencing of human chromosome 16p13.3.
A;Reference number: Z14664
A;Accession: T02345
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1791 <RIC>
A;Cross-references: UNIPROT:O60382; UNIPARC:UPI00000730AE; EMBL:AC004493; NID:g2996648;
C;Genetics:
A;Map position: 16
A;Introns: 1610/2; 1706/2
A;Note: KIAA0324

  Query Match             15.3%;  Score 79.5;  DB 2;  Length 1791;
  Best Local Similarity   34.8%;  Pred. No. 42;
  Matches   24;  Conservative   11;  Mismatches   31;  Indels    3;  Gaps    1;

Qy          35 TKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREK 94
               | |||: ::   : :     | |  ||| | | | ||: | ||::| : ||    |
Db        1615 TPAPKEAVR---EGRPPEPTPAKRKRRSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1671

Qy          95 QQTNSKGSS 103
                 ::|  ||
Db        1672 SSSSSSSSS 1680


RESULT 15
A57640
retinoblastoma binding protein RBQ-1 - human
C;Species: Homo sapiens (man)
C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
C;Accession: A57640
R;Sakai, Y.; Saijo, M.; Coelho, K.; Kishino, T.; Niikawa, N.; Taya, Y.
Genomics 30, 98-101, 1995
A;Title: cDNA sequence and chromosomal localization of a novel human protein, RBQ-1 (RBE
A;Reference number: A57640; MUID:96129310; PMID:8595913
A;Accession: A57640
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-948 <SAK>
A;Cross-references: UNIPROT:Q15290; UNIPARC:UPI000006DB4D; GB:X85133; NID:g728590; PIDN:
C;Genetics:
A;Gene: GDB:RBBP6
A;Cross-references: GDB:626076
A;Map position: 16p12-16p11.2
C;Keywords: tandem repeat; zinc
F;79-129/Domain: RING finger homology <RNG>

  Query Match             15.2%;  Score 79;  DB 2;  Length 948;
  Best Local Similarity   29.8%;  Pred. No. 22;
  Matches   31;  Conservative   17;  Mismatches   48;  Indels    8;  Gaps    4;

Qy           2 SMPDAMPLPGVGEEL-KQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
               ::|     |   ||  :: : ::: ||       :|  :    || :: |: |   |  |
Db         451 TIPTTQAPPLSREEFYREQRRLKEEEKKK-----SKLDEFTNDFAKELMEYKKI-QKERR 504

Qy          61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
               || ||| |  |  || | :||| |     | | :  : |   |:
Db         505 RSFSRSKSPYSGSSY-SRSSYTYSKSRSGSTRSRSYSRSFSRSH 547


Search completed: December 14, 2005, 08:20:57
Job time : 7.39748 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:10:17 ; Search time 33.7918 Seconds
(without alignments)
2171.383 Million cell updates/sec

Title: US-10-683-610-1_COPY_1_104
Perfect score: 521
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....SDDEVSPREKQQTNSKGSSN 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	521	100.0	530	2	Q96PK4_HUMAN	Q96pk4 homo sapien
2	521	100.0	530	2	Q80SW1_MOUSE	Q80sw1 mus musculu
3	521	100.0	607	2	Q502W8_HUMAN	Q502w8 homo sapien
4	521	100.0	623	2	Q6P171_HUMAN	Q6p171 homo sapien
5	454	87.1	540	2	Q5F3Q3_CHICK	Q5f3q3 gallus gall
6	367	70.4	500	1	SAHH2_HUMAN	O43865 homo sapien
7	321.5	61.7	520	2	Q66IP1_XENLA	Q66ip1 xenopus lae
8	295	56.6	611	1	SAHH3_HUMAN	Q96hn2 homo sapien
9	293	56.2	613	2	Q68FL4_MOUSE	Q68fl4 mus musculu
10	288.5	55.4	588	2	Q5BL43_XENTR	Q5bl43 xenopus tro
11	288	55.3	484	2	Q9UG84_HUMAN	Q9ug84 homo sapien
12	283	54.3	512	2	Q7T2B5_BRARE	Q7t2b5 brachydanio
13	282	54.1	483	2	Q5VSM0_HUMAN	Q5vsm0 homo sapien
14	280.5	53.8	588	2	Q68EX1_XENLA	Q68ex1 xenopus lae
15	276	53.0	591	2	Q6PC06_BRARE	Q6pc06 brachydanio
16	275	52.8	508	2	Q8BIH1_MOUSE	Q8bih1 mus musculu
17	273	52.4	508	2	Q5R889_PONPY	Q5r889 pongo pygma
18	270	51.8	491	2	Q4RXQ1_TETNG	Q4rxq1 tetraodon n
19	269	51.6	530	2	Q6ZNS6_HUMAN	Q6zns6 homo sapien
20	257	49.3	541	2	Q4RHJ7_TETNG	Q4rhj7 tetraodon n
21	252	48.4	585	2	Q4RVT5_TETNG	Q4rvt5 tetraodon n
22	250	48.0	583	2	Q6AZU3_XENLA	Q6azu3 xenopus lae
23	159.5	30.6	499	2	Q4S7H8_TETNG	Q4s7h8 tetraodon n
24	126.5	24.3	521	2	Q6NR07_DROME	Q6nr07 drosophila
25	126.5	24.3	521	2	Q9VZX9_DROME	Q9vzx9 drosophila
26	117	22.5	478	2	Q80TQ9_MOUSE	Q80tq9 mus musculu
27	96	18.4	1365	2	Q7PZE8_ANOGA	Q7pze8 anopheles g
28	93	17.9	805	2	Q7RZH4_NEUCR	Q7rzh4 neurospora
29	90	17.3	651	2	Q86B18_DICDI	Q86b18 dictyosteli
30	90	17.3	680	2	Q92033_ANOPU	Q92033 anolis pulc
31	89	17.1	600	2	Q90237_ANOPU	Q90237 anolis pulc
32	88	16.9	447	2	Q6BY19_DEBHA	Q6by19 debaryomyce
33	88	16.9	1266	2	Q80TU6_MOUSE	Q80tu6 mus musculu
34	87.5	16.8	806	2	Q6FSW1_CANGA	Q6fsw1 candida gla
35	87	16.7	345	2	Q7Q7Y6_ANOGA	Q7q7y6 anopheles g
36	87	16.7	478	2	Q61VQ7_CAEBR	Q61vq7 caenorhabdi
37	87	16.7	971	2	Q9XVS4_CAEEL	Q9xvs4 caenorhabdi
38	86.5	16.6	267	2	Q54V71_DICDI	Q54v71 dictyosteli
39	86.5	16.6	414	1	NSR1_YEAST	P27476 saccharomyc
40	86.5	16.6	803	2	Q86J08_DICDI	Q86j08 dictyosteli
41	86.5	16.6	854	2	Q555U5_DICDI	Q555u5 dictyosteli
42	86	16.5	437	2	Q4WJY5_ASPFU	Q4wjy5 aspergillus
43	86	16.5	557	2	Q6KI10_MYCMO	Q6ki10 mycoplasma
44	86	16.5	679	2	Q95XW8_CAEEL	Q95xw8 caenorhabdi
45	85.5	16.4	2223	2	Q5TNJ2_ANOGA	Q5tnj2 anopheles g

ALIGNMENTS

RESULT 1
Q96PK4_HUMAN
ID Q96PK4_HUMAN PRELIMINARY; PRT; 530 AA.
AC Q96PK4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE S-adenosylhomocysteine hydrolase-like protein (S-adenosylhomocysteine
DE hydrolase-like 1).
DE Adenosylhomocysteinase (EC 3.3.1.1) (AdoHcyase) (S-adenosyl-L-
DE homocysteine hydrolase).
GN Name=AHCYL1; ORFNames=RP11-180N18.1-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21901265; PubMed=11904675; DOI=10.1007/s00251-001-0402-z;
RA Dekker J.W., Budhia S., Angel N.Z., Cooper B.J., Clark G.J.,
RA Hart D.N., Kato M.;
RT "Identification of an S-adenosylhomocysteine hydrolase-like transcript
RT induced during dendritic cell differentiation.";
RL Immunogenetics 53:993-1001(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Martin S.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC homocysteine + adenosine.
CC -!- COFACTOR: NAD (By similarity).
CC -!- PATHWAY: Activated methyl cycle.
CC -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR EMBL; AF315687; AAL26869.1; -; mRNA.
DR EMBL; AL772411; CAH70965.1; -; Genomic_DNA.
DR HSSP; P10760; 1B3R.
DR GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR InterPro; IPR000043; Ad_hcy_hydrolase.
DR Pfam; PF05221; AdoHcyase; 1.
DR Pfam; PF00670; AdoHcyase_NAD; 1.
DR TIGRFAMs; TIGR00936; ahcY; 1.
DR PROSITE; PS00738; ADOHCYASE_1; 1.
DR PROSITE; PS00739; ADOHCYASE_2; 1.
KW Hydrolase; NAD; One-carbon metabolism.
SQ SEQUENCE 530 AA; 58951 MW; 974D23361A245D04 CRC64;

Query Match 100.0%; Score 521; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 5.2e-37;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 2
Q80SW1_MOUSE
ID   Q80SW1_MOUSE PRELIMINARY;      PRT;   530 AA.
AC   Q80SW1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   S-adenosylhomocysteine hydrolase-like 1 (IP3R binding protein released
DE   with inositol 1,4,5-trisphosphate) (S-adenosylhomocysteine hydrolase-
DE   like protein).
GN   Name=Ahcyl1; Synonyms=Irbit;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Czech II;
RC   TISSUE=Mammary tumor metastatized to lung. MMTV-LTR/Wnt1 model.
RC   Expression driven by an MMTV-LTR enhancer.;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Czech II;
RC   TISSUE=Mammary tumor metastatized to lung. MMTV-LTR/Wnt1 model.
RC   Expression driven by an MMTV-LTR enhancer.;
RG   NIH MGC Project;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Cerebellum;
RX   MEDLINE=22526701; PubMed=12525476; DOI=10.1074/jbc.M210119200;
RA   Ando H., Mizutani A., Matsu-ura T., Mikoshiba K.;
RT   "IRBIT, a novel inositol 1,4,5-trisphosphate (IP3) receptor-binding
RT   protein, is released from the IP3 receptor upon IP3 binding to the
RT   receptor.";
RL   J. Biol. Chem. 278:10602-10612(2003)..
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21901265; PubMed=11904675; DOI=10.1007/s00251-001-0402-z;
RA   Dekker J.W., Budhia S., Angel N.Z., Cooper B.J., Clark G.J.,
RA   Hart D.N., Kato M.;
RT   "Identification of an S-adenosylhomocysteine hydrolase-like transcript
RT   induced during dendritic cell differentiation.";
RL   Immunogenetics 53:993-1001(2002).
CC   -!- MISCELLANEOUS: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ third party annotation (TPA) entry.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; BC018218; AAH18218.2; -; mRNA.
DR   EMBL; AB092504; BAC65166.1; -; mRNA.
DR   EMBL; BK000547; DAA00059.1; -; mRNA.
DR   HSSP; P10760; 1B3R.
DR   Ensembl; ENSMUSG00000027893; Mus musculus.
DR   MGI; MGI:2385184; Ahcyl1.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
SQ   SEQUENCE   530 AA;  58951 MW;  974D23361A245D04 CRC64;

  Query Match             100.0%;  Score 521;  DB 2;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-37;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104


RESULT 3
Q502W8_HUMAN
ID   Q502W8_HUMAN PRELIMINARY;      PRT;   607 AA.
AC   Q502W8;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
```

```
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RG   NIH MGC Project;
RL   Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; BC095476; AAH95476.1; -; mRNA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR011993; PH_type.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; UNKNOWN_1.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Hydrolase; Hypothetical protein; NAD; One-carbon metabolism.
FT   NON_TER       1      1
SQ   SEQUENCE   607 AA;  66683 MW;  349412833E2D5911 CRC64;

  Query Match             100.0%;  Score 521;  DB 2;  Length 607;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-37;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 137

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db        138 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 181


RESULT 4
Q6P171_HUMAN
ID   Q6P171_HUMAN PRELIMINARY;      PRT;   623 AA.
AC   Q6P171;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Eye;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; BC065254; AAH65254.1; -; mRNA.
DR   HSSP; P10760; 1B3R.
DR   Ensembl; ENSG00000168710; Homo sapiens.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Hydrolase; Hypothetical protein; NAD; One-carbon metabolism.
FT   NON_TER       1      1
SQ   SEQUENCE   623 AA;  68648 MW;  B0A40ED58BAF3F23 CRC64;

  Query Match             100.0%;  Score 521;  DB 2;  Length 623;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-37;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 153

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db        154 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 197


RESULT 5
Q5F3Q3_CHICK
ID   Q5F3Q3_CHICK PRELIMINARY;      PRT;   540 AA.
AC   Q5F3Q3;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=RCJMB04_10b17;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CB; TISSUE=Bursa;
RA   Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA   Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA   Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT   "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT   genefunction analysis.";
RL   Genome Biol. 6:R6-R6(2005).
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; AJ851597; CAH65231.1; -; mRNA.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
```

```
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   540 AA;  59231 MW;  66E479BFA2368F92 CRC64;

  Query Match             87.1%;  Score 454;  DB 2;  Length 540;
  Best Local Similarity   89.4%;  Pred. No. 3.7e-31;
  Matches   93;  Conservative    3;  Mismatches    4;  Indels    4;  Gaps    1;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||:|:    |  ||| ||||| ||||||:|||||||||||||||||||||||||||||
Db          1 MSVPE----PAGGEEAKQAKEAEDAEKFSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 56

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||||||||||||||||||||
Db         57 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 100


RESULT 6
SAHH2_HUMAN
ID   SAHH2_HUMAN     STANDARD;      PRT;   500 AA.
AC   O43865;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Putative adenosylhomocysteinase 2 (EC 3.3.1.1) (S-adenosyl-L-
DE   homocysteine hydrolase) (AdoHcyase).
GN   Name=AHCYL1; Synonyms=XPVKONA;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Skin;
RA   Volpe J.P.G., McDowell M., Jostes R.F., Afzal V., Sadinski W.,
RA   Trask B.J., Legerski R., Cleaver J.E.;
RT   "Complementation of chromosomal instability in the xeroderma
RT   pigmentosum variant by a gene on human chromosome 1 with homology to
RT   S-adenosyl homocysteine hydrolase.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Colon, and Placenta;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: Binds 1 NAD per subunit (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; U82761; AAC01960.1; -; mRNA.
DR   EMBL; BC007576; AAH07576.1; -; mRNA.
DR   EMBL; BC010681; AAH10681.1; -; mRNA.
DR   EMBL; BC016942; AAH16942.1; -; mRNA.
DR   HSSP; P10760; 1B3R.
DR   Ensembl; ENSG00000168710; Homo sapiens.
DR   HGNC; HGNC:344; AHCYL1.
DR   MIM; 607826; -.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   PANTHER; PTHR11784; Ad_hcy_hydrolase; 1.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
FT   REGION      251    418       NAD binding (By similarity).
FT   BINDING     125    125       Substrate (By similarity).
FT   BINDING     199    199       Substrate (By similarity).
FT   BINDING     224    224       Substrate (By similarity).
FT   BINDING     254    254       Substrate (By similarity).
FT   BINDING     258    258       Substrate (By similarity).
SQ   SEQUENCE   500 AA;  55628 MW;  FDEABFC70044B9DA CRC64;

  Query Match             70.4%;  Score 367;  DB 1;  Length 500;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-23;
  Matches   74;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         31 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 60

Qy         91 PREKQQTNSKGSSN 104
              ||||||||||||||
Db         61 PREKQQTNSKGSSN 74


RESULT 7
Q66IP1_XENLA
ID   Q66IP1_XENLA PRELIMINARY;      PRT;   520 AA.
AC   Q66IP1;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC86404 protein.
GN   Name=MGC86404;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
```

```
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RA   Klein S., Gerhard D.S.;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC081269; AAH81269.1; -; mRNA.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
SQ   SEQUENCE   520 AA;  57839 MW;  C73583CF180234C3 CRC64;

  Query Match             61.7%;  Score 321.5;  DB 2;  Length 520;
  Best Local Similarity   67.3%;  Pred. No. 1.2e-19;
  Matches   70;  Conservative    9;  Mismatches   14;  Indels   11;  Gaps    2;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||:|:|:              |   |||: ||| |:| ||| ||||||||||:|| |||||
Db          1 MSLPEAV----------TTDEPAAAEKFLFMAPVSKTPKK-IQFADDMQEFSKFRTKTGR 49

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||| ||||||||||:|||:|||  |||| :
Db         50 RSLSRSISQSSTDSYSSVASYTDSSDDEMSPRDKQQITSKGSGD 93


RESULT 8
SAHH3_HUMAN
ID   SAHH3_HUMAN    STANDARD;      PRT;   611 AA.
AC   Q96HN2; O94917;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Putative adenosylhomocysteinase 3 (EC 3.3.1.1) (S-adenosyl-L-
DE   homocysteine hydrolase) (AdoHcyase).
GN   Name=KIAA0828;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Brain;
RX   MEDLINE=99156230; PubMed=10048485;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 5:355-364(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Lung, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: Binds 1 NAD per subunit (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AB020635; BAA74851.1; ALT_INIT; mRNA.
DR   EMBL; BC008349; AAH08349.1; -; mRNA.
DR   EMBL; BC024325; AAH24325.1; -; mRNA.
DR   HSSP; P23526; 1LI4.
DR   Ensembl; ENSG00000158467; Homo sapiens.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   PANTHER; PTHR11784; Ad_hcy_hydrolase; 1.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
FT   REGION      362    529       NAD binding (By similarity).
FT   BINDING     236    236       Substrate (By similarity).
FT   BINDING     310    310       Substrate (By similarity).
FT   BINDING     335    335       Substrate (By similarity).
FT   BINDING     365    365       Substrate (By similarity).
FT   BINDING     369    369       Substrate (By similarity).
FT   CONFLICT    122    122       Missing (in Ref. 1).
SQ   SEQUENCE   611 AA;  66721 MW;  A76B1275FC0D891D CRC64;

  Query Match             56.6%;  Score 295;  DB 1;  Length 611;
  Best Local Similarity   55.2%;  Pred. No. 3e-17;
  Matches   69;  Conservative   11;  Mismatches   17;  Indels   28;  Gaps    3;

Qy          8 PLPGVGE----------ELKQAKEIEDAEKY-------------SFMATVTKAP---K 39
              |:|| |             :: ||  :: :: :             |   |||:||   |
Db         61 PVPGPGSGPAAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 120

Qy         40 KQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNS 99
              |||||||  ||| | ||| ||||||||||||||||||||||||||||| |||:||| ||
Db        121 KQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNS 180

Qy        100 KGSSN 104
              ||||:
```

```
Db        181 KGSSD 185


RESULT 9
Q68FL4_MOUSE
ID   Q68FL4_MOUSE PRELIMINARY;      PRT;   613 AA.
AC   Q68FL4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   RIKEN cDNA 4631427C17.
GN   Name=4631427C17Rik;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Director MGC Project;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; BC079660; AAH79660.1; -; mRNA.
DR   Ensembl; ENSMUSG00000029772; Mus musculus.
DR   MGI; MGI:1921590; 4631427C17Rik.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
SQ   SEQUENCE   613 AA;  66899 MW;  4413DBF1F4825860 CRC64;

  Query Match             56.2%;  Score 293;  DB 2;  Length 613;
  Best Local Similarity   55.2%;  Pred. No. 4.5e-17;
  Matches   69;  Conservative   10;  Mismatches   18;  Indels   28;  Gaps    3;

Qy          8 PLPGVGE-----------ELKQAKEIEDAEKY-------------SFMATVTKAP---K 39
              | || |             :: ||  :: :: :             |   |||:||   |
Db         63 PAPGPGSGPTAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 122

Qy         40 KQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNS 99
              |||||||  ||| | ||| |||||||||||||||||||||||||||||||| |||:||| ||
Db        123 KQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNS 182

Qy        100 KGSSN 104
              ||||:
Db        183 KGSSD 187


RESULT 10
Q5BL43_XENTR
ID   Q5BL43_XENTR PRELIMINARY;      PRT;   588 AA.
AC   Q5BL43;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein LOC549163.
GN   Name=LOC549163;
OS   Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Silurana.
OX   NCBI_TaxID=8364;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RA   Klein S., Gerhard D.S.;
RL   Submitted (FEB-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC090609; AAH90609.1; -; mRNA.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   InterPro; IPR007698; AlaDh_PNT_C.
DR   InterPro; IPR001379; Egg_lysin.
DR   InterPro; IPR001093; IMPDh/GMPRtase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; UNKNOWN_1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   588 AA;  64721 MW;  2CBAD9A46134581C CRC64;

  Query Match             55.4%;  Score 288.5;  DB 2;  Length 588;
  Best Local Similarity   57.3%;  Pred. No. 1.1e-16;
  Matches   67;  Conservative   12;  Mismatches   21;  Indels   17;  Gaps    3;
```

```
Qy          4 PDAMPLPGVGEELKQA----------KEIEDAEKYSFMATVTKAPK--KQIQFADD 47
              ||  |:| : : :| |            |:  | |  |    |::||:  |:|||||
Db         47 PDNSPIPALSQPIKGAPASAMKRSDPTQRPKDGGD-EAGSPEGAVSEAPRTVKKIQFADQ 105

Qy         48 MQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
                :|| | ||| |||||||||||||||||||||||||||||| |||:||| ||||||:
Db        106 KREFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSKGSSD 162


RESULT 11
Q9UG84_HUMAN
ID   Q9UG84_HUMAN PRELIMINARY;      PRT;   484 AA.
AC   Q9UG84;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Adenosylhomocysteinase (EC 3.3.1.1) (AdoHcyase) (S-adenosyl-L-
DE   homocysteine hydrolase) (Fragment).
GN   Name=DKFZp564A1523;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; AL049954; CAB43223.2; -; mRNA.
DR   PIR; T08681; T08681.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; Hypothetical protein; NAD; One-carbon metabolism.
FT   NON_TER       1       1
SQ   SEQUENCE   484 AA;  53868 MW;  8A55EF564F6B7E1D CRC64;

  Query Match             55.3%;  Score 288;  DB 2;  Length 484;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-17;
  Matches   58;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         47 DMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 58


RESULT 12
Q7T2B5_BRARE
ID   Q7T2B5_BRARE PRELIMINARY;      PRT;   512 AA.
AC   Q7T2B5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   S-adenosylhomocysteine hydrolase-like 1.
GN   Name=ahcyl1;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC054614; AAH54614.1; -; mRNA.
DR   HSSP; P10760; 1B3R.
DR   Ensembl; ENSDARG00000001500; Danio rerio.
DR   ZFIN; ZDB-GENE-021206-7; ahcyl1.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase.
SQ   SEQUENCE   512 AA;  56775 MW;  3E0A09DFFE086683 CRC64;

  Query Match             54.3%;  Score 283;  DB 2;  Length 512;
  Best Local Similarity   76.9%;  Pred. No. 2.7e-16;
  Matches   60;  Conservative    5;  Mismatches   11;  Indels    2;  Gaps    1;

Qy         29 SFMATVTKAPKKQIQFA--DDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD 86
              :| |  |    |:||||  :| |||:||||||||||||||||||||||||||||||||||
Db          9 AFTAVPTHKSVKKIQFANQEDKQEFSKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD 68

Qy         87 DEVSPREKQQTNSKGSSN 104
              || ||||| | ||||||:
Db         69 DETSPREKSQVNSKGSSD 86


RESULT 13
Q5VSM0_HUMAN
ID   Q5VSM0_HUMAN PRELIMINARY;      PRT;   483 AA.
AC   Q5VSM0;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Adenosylhomocysteinase (EC 3.3.1.1) (AdoHcyase) (S-adenosyl-L-
DE   homocysteine hydrolase).
GN   Name=AHCYL1; ORFNames=RP11-180N18.1-002;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
```

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Martin S.;
RL   Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = L-
CC       homocysteine + adenosine.
CC   -!- COFACTOR: NAD (By similarity).
CC   -!- PATHWAY: Activated methyl cycle.
CC   -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
DR   EMBL; AL772411; CAH70966.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000168710; Homo sapiens.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase; NAD; One-carbon metabolism.
SQ   SEQUENCE   483 AA;  53753 MW;  18DFC7E74697E1D4 CRC64;

  Query Match             54.1%;  Score 282;  DB 2;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-16;
  Matches   57;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         48 MQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 57


RESULT 14
Q68EX1_XENLA
ID   Q68EX1_XENLA PRELIMINARY;      PRT;   588 AA.
AC   Q68EX1;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   MGC84148 protein.
GN   Name=MGC84148;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RA   Klein S., Gerhard D.S.;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC080079; AAH80079.1; -; mRNA.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
SQ   SEQUENCE   588 AA;  64853 MW;  81A3869C48D026B8 CRC64;

  Query Match             53.8%;  Score 280.5;  DB 2;  Length 588;
  Best Local Similarity   56.4%;  Pred. No. 5.2e-16;
  Matches   66;  Conservative   12;  Mismatches   22;  Indels   17;  Gaps    3;

Qy          4 PDAMPLPGVGEELKQA----------KEIEDAEKYSFMATVTKAPK--KQIQFADD 47
              |:  |:| :   :| |          |:  | |  |    |::||:  |:|||||
Db         47 PEISPIPALSLPMKGAPASAMKRTDPTLRPKDGGD-EAGSPEGAVSEAPRTVKKIQFADQ 105

Qy         48 MQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
                :|| | ||| ||||||||||||||||||||||||||||| |||:||| ||||||:
Db        106 KREFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSKGSSD 162


RESULT 15
Q6PC06_BRARE
ID   Q6PC06_BRARE PRELIMINARY;      PRT;   591 AA.
AC   Q6PC06;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   S-adenosylhomocysteine hydrolase-like 2.
GN   Name=ahcyl2;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Wild-type; TISSUE=Eye;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
```

```
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Wild-type; TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC059517; AAH59517.1; -; mRNA.
DR   HSSP; P10760; 1B3R.
DR   ZFIN; ZDB-GENE-040115-5; ahcyl2.
DR   GO; GO:0004013; F:adenosylhomocysteinase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR   InterPro; IPR000043; Ad_hcy_hydrolase.
DR   Pfam; PF05221; AdoHcyase; 1.
DR   Pfam; PF00670; AdoHcyase_NAD; 1.
DR   TIGRFAMs; TIGR00936; ahcY; 1.
DR   PROSITE; PS00738; ADOHCYASE_1; 1.
DR   PROSITE; PS00739; ADOHCYASE_2; 1.
KW   Hydrolase.
SQ   SEQUENCE   591 AA;  65087 MW;  6D916B5342EA7D15 CRC64;

  Query Match             53.0%;  Score 276;  DB 2;  Length 591;
  Best Local Similarity   58.1%;  Pred. No. 1.3e-15;
  Matches   61;  Conservative   10;  Mismatches   28;  Indels    6;  Gaps    2;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPK-KQIQFADDMQEFTKFPTKTG 59
              | || |  :     :    :   : :      || : |: |:|||||  ||| | ||| |
Db         66 MKMPQASAMKRTDPQQNGGEAFVNCD-----GTVAEGPRMKKIQFADQKQEFNKRPTKIG 120

Qy         60 RRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSN 104
              |||||||||||||||||||||||||||| |||:||| ||||| :
Db        121 RRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSKGSGD 165


Search completed: December 14, 2005, 08:20:13
Job time : 34.7918 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2005, 08:09:37 ; Search time 132.918 Seconds
(without alignments)
1751.988 Million cell updates/sec

Title: US-10-683-610-1
Perfect score: 2781
Sequence: 1 MSMPDAMPLPGVGEELKQAK.....QAKYLGLNKNGPFKPNYYRY 530

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2781	100.0	530	7	ADB68448	Adb68448 Human DCA
2	2781	100.0	530	7	ADB68450	Adb68450 Murine DC
3	2781	100.0	530	7	ADB68446	Adb68446 Human DCA
4	2781	100.0	530	8	ADM94432	Adm94432 Mouse ino
5	2781	100.0	530	8	ABM80118	Abm80118 Tumour-as
6	2781	100.0	530	8	ADP23045	Adp23045 PRO polyp
7	2781	100.0	614	2	AAW56097	Aaw56097 Amino aci
8	2778	99.9	530	7	ADB68452	Adb68452 Murine DC
9	2618	94.1	500	2	AAW90061	Aaw90061 Human SAH
10	2582	92.8	597	8	ADO43989	Ado43989 Amino aci
11	2456	88.3	611	8	ABM80935	Abm80935 Tumour-as
12	2429	87.3	508	8	ADS10638	Ads10638 Human the
13	2429	87.3	509	8	ADL90193	Adl90193 Human enz
14	2429	87.3	564	5	ABG32371	Abg32371 Novel hum
15	1970	70.8	521	4	ABB58663	Abb58663 Drosophil
16	1819.5	65.4	504	4	ABB71710	Abb71710 Drosophil
17	1601.5	57.6	455	4	ABG12029	Abg12029 Novel hum
18	1513	54.4	317	8	ADR10475	Adr10475 Human pro
19	1307.5	47.0	441	9	AEB39552	Aeb39552 L. pneumo
20	1307.5	47.0	447	9	AEB36129	Aeb36129 L. pneumo
21	1289	46.4	426	8	ADS21102	Ads21102 Bacterial
22	1187	42.7	461	8	ADN24890	Adn24890 Bacterial
23	1187	42.7	461	8	ADN22130	Adn22130 Bacterial
24	1185	42.6	463	8	ADS30671	Ads30671 Bacterial
25	1158	41.6	466	8	ADS23077	Ads23077 Bacterial
26	1154.5	41.5	432	4	ABB70895	Abb70895 Drosophil
27	1154.5	41.5	468	9	AEA20951	Aea20951 Novel hum
28	1151.5	41.4	457	3	AAB43614	Aab43614 Human can
29	1151	41.4	432	7	ADF44376	Adf44376 Human pla
30	1151	41.4	432	7	ADF76060	Adf76060 Wild type
31	1146	41.2	432	8	ABM82227	Abm82227 Tumour-as
32	1146	41.2	432	8	ADS88176	Ads88176 Human pro
33	1146	41.2	432	8	ADP23057	Adp23057 PRO polyp
34	1146	41.2	432	9	ADX07800	Adx07800 Cyclin-de
35	1146	41.2	432	9	ADX56955	Adx56955 Human S-a
36	1146	41.2	432	9	ADX57572	Adx57572 Human S-a
37	1145.5	41.2	433	5	AAE25380	Aae25380 Human NZM
38	1142	41.1	461	8	ADS25619	Ads25619 Bacterial
39	1142	41.1	461	8	ADS22433	Ads22433 Bacterial
40	1142	41.1	461	8	ADS42602	Ads42602 Bacterial
41	1142	41.1	464	8	ADS26092	Ads26092 Bacterial
42	1141	41.0	437	8	ADN22444	Adn22444 Bacterial
43	1130.5	40.7	510	3	AAG34059	Aag34059 Zea mays
44	1129	40.6	433	8	ADN19777	Adn19777 Bacterial
45	1129	40.6	433	8	ADN19772	Adn19772 Bacterial

ALIGNMENTS

RESULT 1
ADB68448
ID ADB68448 standard; protein; 530 AA.
XX
AC ADB68448;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human DCAL protein - SEQ ID 4.
XX
KW dendritic cell expressed S-adenosyl homocysteine hydrolase-like molecule;
KW DC; AHCY; DCAL; antiallergic; immunosuppressive; allergy; asthma;
KW allergic rhinitis; systemic anaphylaxis; autoimmune; diabetes mellitus;
KW rheumatoid arthritis; transplant rejection; vaccine; gene therapy; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1. .106
FT /note = Hydrophilic domain
FT Domain 107. .530
FT /note = adenosine-homocysteine hydrolase (AHCY)-like
FT domain
FT Binding-site 318. .323
FT /note= "NAD+ binding motif (GXGXXG)"
XX
PN WO2003055997-A1.
XX
PD 10-JUL-2003.
XX
PF 24-DEC-2002; 2002WO-AU001761.
XX
PR 24-DEC-2001; 2001AU-00009741.
XX
PA (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.
XX
PI Kato M, Angel NZ, Cooper BJ, Hart DNJ;
XX
DR WPI; 2003-559275/52.
DR N-PSDB; ADB68445, ADB68447.
XX
PT New dendritic cell expressed S-adenosyl homocysteine hydrolase-like
PT molecule (DCAL) gene, useful for preparing a composition for treating or
PT preventing a condition e.g., allergy, autoimmune disease or transplant
PT rejection.
XX
PS Disclosure; Fig 1B; 189pp; English.

```
XX
CC   The invention relates to a novel isolated polynucleotide comprising a
CC   dendritic cell (DC)-expressed S-adenosyl homocysteine hydrolase (AHCY)-
CC   like molecule (DCAL) gene. The polynucleotide of the invention
CC   demonstrates antiallergic and immunosuppressive activities and may be
CC   useful for treating a condition including an allergy such as asthma,
CC   allergic rhinitis or systemic anaphylaxis, an autoimmune disease such as
CC   diabetes mellitus or rheumatoid arthritis or transplant rejection.
CC   Furthermore, the polynucleotide may be useful as a vaccine or during gene
CC   therapy procedures. The genetically modified animal or the identified
CC   modulatory agent of the invention may be used in the study of immunity,
CC   DC function, brain physiology or neuronal cell function. The current
CC   sequence is that of the human DCAL protein (SEQ ID 4) of the invention
CC   which is encoded by the human DCAL cDNA.
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 2781;  DB 7;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-268;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 2
ADB68450
ID   ADB68450 standard; protein; 530 AA.
XX
AC   ADB68450;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Murine DCAL protein - SEQ ID 6.
XX
KW   dendritic cell expressed S-adenosyl homocysteine hydrolase-like molecule;
KW   DC; AHCY; DCAL; antiallergic; immunosuppressive; allergy; asthma;
KW   allergic rhinitis; systemic anaphylaxis; autoimmune; diabetes mellitus;
KW   rheumatoid arthritis; transplant rejection; vaccine; gene therapy;
KW   murine; mouse.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Binding-site    318. .323
FT                   /note= "NAD+ binding motif (GXGXXG)"
XX
PN   WO2003055997-A1.
XX
PD   10-JUL-2003.
XX
PF   24-DEC-2002; 2002WO-AU001761.
XX
PR   24-DEC-2001; 2001AU-00009741.
XX
PA   (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.
XX
PI   Kato M,  Angel NZ,  Cooper BJ,  Hart DNJ;
XX
DR   WPI; 2003-559275/52.
DR   N-PSDB; ADB68449.
XX
PT   New dendritic cell expressed S-adenosyl homocysteine hydrolase-like
PT   molecule (DCAL) gene, useful for preparing a composition for treating or
PT   preventing a condition e.g., allergy, autoimmune disease or transplant
PT   rejection.
XX
PS   Disclosure; Page 170-172; 189pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising a
CC   dendritic cell (DC)-expressed S-adenosyl homocysteine hydrolase (AHCY)-
CC   like molecule (DCAL) gene. The polynucleotide of the invention
CC   demonstrates antiallergic and immunosuppressive activities and may be
CC   useful for treating a condition including an allergy such as asthma,
CC   allergic rhinitis or systemic anaphylaxis, an autoimmune disease such as
CC   diabetes mellitus or rheumatoid arthritis or transplant rejection.
CC   Furthermore, the polynucleotide may be useful as a vaccine or during gene
CC   therapy procedures. The genetically modified animal or the identified
CC   modulatory agent of the invention may be used in the study of immunity,
CC   DC function, brain physiology or neuronal cell function. The current
CC   sequence is that of the murine DCAL protein (SEQ ID 6) of the invention
CC   which is encoded by genomic DNA.
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 2781;  DB 7;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-268;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
```

```
Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 3
ADB68446
ID   ADB68446 standard; protein; 530 AA.
XX
AC   ADB68446;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human DCAL protein - SEQ ID 2.
XX
KW   dendritic cell expressed S-adenosyl homocysteine hydrolase-like molecule;
KW   DC; AHCY; DCAL; antiallergic; immunosuppressive; allergy; asthma;
KW   allergic rhinitis; systemic anaphylaxis; autoimmune; diabetes mellitus;
KW   rheumatoid arthritis; transplant rejection; vaccine; gene therapy; human;
KW   chromosome 1p21.1.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .106
FT                   /note = Hydrophilic domain
FT   Domain          107. .530
FT                   /note = adenosine-homocysteine hydrolase (AHCY)-like
FT                   domain
FT   Binding-site    318. .323
FT                   /note= "NAD+ binding motif (GXGXXG)"
XX
PN   WO2003055997-A1.
XX
PD   10-JUL-2003.
XX
PF   24-DEC-2002; 2002WO-AU001761.
XX
PR   24-DEC-2001; 2001AU-00009741.
XX
PA   (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.
XX
PI   Kato M,  Angel NZ,  Cooper BJ,  Hart DNJ;
XX
DR   WPI; 2003-559275/52.
DR   N-PSDB; ADB68445, ADB68447.
XX
PT   New dendritic cell expressed S-adenosyl homocysteine hydrolase-like
PT   molecule (DCAL) gene, useful for preparing a composition for treating or
PT   preventing a condition e.g., allergy, autoimmune disease or transplant
PT   rejection.
XX
PS   Disclosure; Fig 1B; 189pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising a
CC   dendritic cell (DC)-expressed S-adenosyl homocysteine hydrolase (AHCY)-
CC   like molecule (DCAL) gene. The polynucleotide of the invention
CC   demonstrates antiallergic and immunosuppressive activities and may be
CC   useful for treating a condition including an allergy such as asthma,
CC   allergic rhinitis or systemic anaphylaxis, an autoimmune disease such as
CC   diabetes mellitus or rheumatoid arthritis or transplant rejection.
CC   Furthermore, the polynucleotide may be useful as a vaccine or during gene
CC   therapy procedures. The genetically modified animal or the identified
CC   modulatory agent of the invention may be used in the study of immunity,
CC   DC function, brain physiology or neuronal cell function. The current
CC   sequence is that of the human DCAL protein (SEQ ID 2) of the invention
CC   which is encoded by DNA localised to chromosome 1p21.1.
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 2781;  DB 7;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-268;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 4
ADM94432
ID   ADM94432 standard; protein; 530 AA.
XX
AC   ADM94432;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Mouse inositol 1,4,5-triphosphate receptor binding protein.
XX
KW   inositol-triphosphate receptor-binding protein; IP3 receptor;
KW   fluorescence resonance energy transfer; IP3 quantification;
KW   IP3-induced Ca2+ releasing activity; PIP2 function.
XX
OS   Mus musculus.
XX
PN   EP1408049-A2.
XX
PD   14-APR-2004.
XX
PF   13-OCT-2003; 2003EP-00256457.
XX
PR   11-OCT-2002; 2002JP-00299429.
XX
PA   (RIKE ) RIKEN KK.
XX
PI   Mikoshiba K,  Ando H,  Mizutani A,  Matsu-Ura T;
```

XX
DR WPI; 2004-349314/33.
DR N-PSDB; ADM94433.
XX
PT New inositol 1,4,5-trisphosphate receptor-binding protein that binds to
PT inositol 1,4,5-trisphosphate receptor, useful as inositol 1,4,5-
PT trisphosphate indicator intracellularly and in cell-free system.
XX
PS Claim 1; SEQ ID NO 1; 61pp; English.
XX
CC The invention relates to a novel inositol 1,4,5-triphosphate (IP3)
CC receptor-binding protein (I) that binds to IP3 receptor. (I) is used as
CC an IP3 indicator in combination with a fluorescent molecule for detecting
CC IP3 by the fluorescence resonance energy transfer (FRET) method using a
CC combination of two fluorescent molecules applicable to the FRET method,
CC and a protein containing at least IP3-binding domain of IP3 receptor
CC labelled with the other fluorescent molecule. (I) is useful as IP3
CC indicator intracellularly and in cell-free system, for detecting and/or
CC quantifying IP3, as modulator of IP3-induced Ca2+ releasing activity,
CC modulator of PIP2 function and Na2+ and bicarbonate ions. This sequence
CC represents the IP3 receptor binding protein.
XX
SQ Sequence 530 AA;

Query Match 100.0%; Score 2781; DB 8; Length 530;
Best Local Similarity 100.0%; Pred. No. 8.7e-268;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
Db 1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy 61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
Db 61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy 121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
Db 121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy 181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
Db 181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy 241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
Db 241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy 301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
Db 301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy 361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
Db 361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy 421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
Db 421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy 481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
Db 481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530

RESULT 5
ABM80118
ID ABM80118 standard; protein; 530 AA.
XX
AC ABM80118;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO80619, SEQ:301.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN37426.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 12; SEQ ID NO 301; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 530 AA;

Query Match 100.0%; Score 2781; DB 8; Length 530;
Best Local Similarity 100.0%; Pred. No. 8.7e-268;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
Db 1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy 61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
Db 61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy 121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
Db 121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

```
Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 6
ADP23045
ID   ADP23045 standard; protein; 530 AA.
XX
AC   ADP23045;
XX
DT   18-NOV-2004  (first entry)
XX
DE   PRO polypeptide SEQ ID NO:139.
XX
KW   PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW   osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW   antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX
OS   Unidentified.
XX
PN   WO2004041170-A2.
XX
PD   21-MAY-2004.
XX
PF   30-OCT-2003; 2003WO-US034312.
XX
PR   01-NOV-2002; 2002US-0423394P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Clark H,  Schoenfeld J,  Van Lookeren M,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-419628/39.
DR   N-PSDB; ADP23044.
XX
PT   New PRO polypeptides and polynucleotides, useful for treating e.g.
PT   erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT   renal disease, or demyelinating diseases of the central or peripheral
PT   nervous system.
XX
PS   Claim 7; SEQ ID NO 139; 2940pp; English.
XX
CC   The invention relates to a novel isolated nucleic acid and the PRO
CC   polypeptide encoded by it. A protein of the invention has
CC   antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC   osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC   antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC   of the invention may have a use in gene therapy. The PRO polypeptide, its
CC   agonist, antagonist, or antibody that specifically binds to the
CC   polypeptide is useful for treating an immune related disorder such as
CC   systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC   juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC   idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC   vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC   thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC   disease, a demyelinating disease of the central or peripheral nervous
CC   system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC   a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC   disease, infectious or autoimmune chronic active hepatitis, primary
CC   biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC   inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC   disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC   disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC   disease, asthma, allergic rhinitis, atopic dermatitis, food
CC   hypersensitivity, urticaria, an immunologic disease of the lung,
CC   eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC   pneumonitis, a transplantation associated disease, graft rejection or
CC   graft-versus-host disease. The present sequence represents a PRO protein
CC   of the invention.
XX
SQ   Sequence 530 AA;

  Query Match             100.0%;  Score 2781;  DB 8;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-268;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 7
AAW56097
ID   AAW56097 standard; protein; 614 AA.
XX
AC   AAW56097;
XX
DT   21-AUG-1998  (first entry)
XX
DE   Amino acid sequence of the 0DD4b5.3 enzyme.
XX
KW   DD4b9.3 enzyme; S-adenosyl-L-homocysteine hydrolase-type activity;
```

```
KW   AHCY-type activity; stimulation; inhibition; dendritic cell; cancer;
KW   autoimmune disease; transplantation.
XX
OS   Homo sapiens.
XX
PN   WO9814562-A1.
XX
PD   09-APR-1998.
XX
PF   06-OCT-1997;   97WO-NZ000133.
XX
PR   04-OCT-1996;   96NZ-00299507.
XX
PA   (HART/) HART D N J.
XX
PI   Hart DNJ;
XX
DR   WPI; 1998-240074/21.
DR   N-PSDB; AAV28617.
XX
PT   New isolated S-adenosyl-L-homocysteine hydrolase enzyme - is used to
PT   develop products which can be used in the treatment of e.g. auto-immune
PT   disease, transplantations or cancers.
XX
PS   Claim 3; Fig 1; 33pp; English.
XX
CC   This is the amino acid sequence of the DD4b9.3 enzyme which has S-
CC   adenosyl-L-homocysteine hydrolase (AHCY)-type activity. Its products can
CC   be used to identify substances which have a stimulatory or inhibitory
CC   effect on the enzyme activity. Such substances can be used to modulate
CC   dendritic cell (DC) function and for immunomodulation. They can be used
CC   in the treatment of e.g. autoimmune diseases, transplantations or
CC   cancers. The products can also be used for detection and diagnosis
XX
SQ   Sequence 614 AA;

  Query Match             100.0%;  Score 2781;  DB 2;  Length 614;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-267;
  Matches  530;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 144

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 204

Qy        121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 264

Qy        181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 324

Qy        241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 384

Qy        301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 444

Qy        361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 504

Qy        421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 564

Qy        481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 614


RESULT 8
ADB68452
ID   ADB68452 standard; protein; 530 AA.
XX
AC   ADB68452;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Murine DCAL protein - SEQ ID 8.
XX
KW   dendritic cell expressed S-adenosyl homocysteine hydrolase-like molecule;
KW   DC; AHCY; DCAL; antiallergic; immunosuppressive; allergy; asthma;
KW   allergic rhinitis; systemic anaphylaxis; autoimmune; diabetes mellitus;
KW   rheumatoid arthritis; transplant rejection; vaccine; gene therapy;
KW   murine; mouse.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Binding-site    318. .323
FT                   /note= "NAD+ binding motif (GXGXXG)"
XX
PN   WO2003055997-A1.
XX
PD   10-JUL-2003.
XX
PF   24-DEC-2002; 2002WO-AU001761.
XX
PR   24-DEC-2001; 2001AU-00009741.
XX
PA   (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.
XX
PI   Kato M,  Angel NZ,  Cooper BJ,  Hart DNJ;
XX
DR   WPI; 2003-559275/52.
DR   N-PSDB; ADB68451.
XX
PT   New dendritic cell expressed S-adenosyl homocysteine hydrolase-like
PT   molecule (DCAL) gene, useful for preparing a composition for treating or
PT   preventing a condition e.g., allergy, autoimmune disease or transplant
PT   rejection.
XX
PS   Disclosure; Page 175-177; 189pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising a
CC   dendritic cell (DC)-expressed S-adenosyl homocysteine hydrolase (AHCY)-
CC   like molecule (DCAL) gene. The polynucleotide of the invention
CC   demonstrates antiallergic and immunosuppressive activities and may be
CC   useful for treating a condition including an allergy such as asthma,
CC   allergic rhinitis or systemic anaphylaxis, an autoimmune disease such as
CC   diabetes mellitus or rheumatoid arthritis or transplant rejection.
CC   Furthermore, the polynucleotide may be useful as a vaccine or during gene
CC   therapy procedures. The genetically modified animal or the identified
CC   modulatory agent of the invention may be used in the study of immunity,
CC   DC function, brain physiology or neuronal cell function. The current
CC   sequence is that of the murine DCAL protein (SEQ ID 8) of the invention
CC   which is encoded by cDNA.
XX
SQ   Sequence 530 AA;

  Query Match             99.9%;  Score 2778;  DB 7;  Length 530;
  Best Local Similarity   99.8%;  Pred. No. 1.7e-267;
  Matches  529;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSMPDAMPLPGVGEELKQAKEIEDAEKYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGR 60

Qy         61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 RSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREI 120

Qy         121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 EIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCRWSACNIY 180

Qy         181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 STQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKK 240

Qy         241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 YPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDG 300

Qy         301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 LKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMDGFRVVKL 360

Qy         361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 NEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVR 420

Qy         421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 SQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAPEGRYKQD 480

Qy         481 VYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
                |||||||||||||||||:||||||||||||||||||||||||||||||
Db         481 VYLLPKKMDEYVASLHLPTFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530


RESULT 9
AAW90061
ID   AAW90061 standard; protein; 500 AA.
XX
AC   AAW90061;
XX
DT   04-MAR-1999  (first entry)
XX
DE   Human SAHH protein.
XX
KW   S-adenosyl-5-homocysteine hydrolase; SAHH; human; drug screening;
KW   treatment; infection; cancer; autoimmune disease; detection; diagnosis;
KW   gene mapping; antisense; therapy; antagonist; immunoassay.
XX
OS   Homo sapiens.
XX
PN   US5854023-A.
XX
PD   29-DEC-1998.
XX
PF   17-JUL-1997;   97US-00896005.
XX
PR   17-JUL-1997;   97US-00896005.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Shah P,  Lal P,  Hillman JL,  Corley NC;
XX
DR   WPI; 1999-094906/08.
DR   N-PSDB; AAV73924.
XX
PT   Nucleic acid encoding human S-adenosyl-5-homocysteine hydrolase - for
PT   production of recombinant enzyme, useful for diagnosis, treatment and
PT   prevention of cancers, infections and autoimmune diseases.
XX
PS   Claim 1; Fig 1A-I; 40pp; English.
XX
CC   This sequence represents a human S-adenosyl-5-homocysteine hydrolase
CC   (SAHH). The SAHH protein can be used to generate specific antibodies and
CC   in drug screening to identify specific binding agents. Antagonists of the
CC   protein are used to treat or prevent a wide range of viral, bacterial,
CC   fungal, parasitic, protozoal or helminthic infections, many cancers
CC   (leukaemia, lymphoma or solid tumours), and many autoimmune diseases
CC   (e.g. acquired immune deficiency syndrome, allergy, asthma, diabetes
CC   mellitus, multiple sclerosis etc). All these conditions may be treated by
CC   expressing antisense sequences, triplex-forming agents or ribozymes
CC   directed against the nucleic acid. The nucleic acid and its fragments can
CC   be used as probes or primers for detecting and quantifying gene
CC   expression, for diagnosis or monitoring of disease, to identify genetic
CC   variations, mutations or polymorphisms, in gene mapping and as antisense
CC   therapeutics. Antibodies are used directly as antagonists, indirectly to
CC   deliver active agents to SAHH-expressing cells, to diagnose and monitor
CC   diseases in standard immunoassays, in competitive drug screens and to
CC   isolate the protein from natural sources
XX
SQ   Sequence 500 AA;

  Query Match            94.1%;  Score 2618;  DB 2;  Length 500;
  Best Local Similarity  99.8%;  Pred. No. 1.5e-251;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          31 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 90
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVS 60

Qy          91 PREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVG 150
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 PREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVG 120

Qy         151 CTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWW 210
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWW 180

Qy         211 CIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGK 270
               |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         181 CIDRCVNMDGWQANMILDDGGDLTHWVCKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGK 240

Qy         271 LCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAAL 330
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 LCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAAL 300

Qy         331 KALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKN 390
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 KALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKN 360

Qy         391 SCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTV 450
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 SCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTV 420

Qy         451 PTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDD 510
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 PTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDD 480

Qy         511 QAKYLGLNKNGPFKPNYYRY 530
               ||||||||||||||||||||
Db         481 QAKYLGLNKNGPFKPNYYRY 500


RESULT 10
ADO43989
ID   ADO43989 standard; protein; 597 AA.
XX
AC   ADO43989;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Amino acid sequence of human putative S-adenosylhomocysteinase.
XX
KW   protein complex; neurological disease; stroke; neurodegeneration;
KW   Wallerian degeneration; Alzheimer's disease; neurological disorder;
KW   epilepsy; inflammatory condition; ulcerative colitis; Crohn's disease;
KW   atherosclerosis; 1D-MYO-inositol triphosphate 3 kinase A; ASK1; ASK2;
KW   ASK3; CaMKII beta; CaMKII delta; CaMKII gamma; casein kinase II alpha;
KW   Cdc37; CHK2; CTCL tumour antigen SE20-4; EF-1 alpha 1; EMAP;
```

```
KW   FLJ14653 NT2RP2002252; FLJ30839 FEBRA2002429; HERC2;
KW   inositol polyphosphate-5-phosphatase; inositol-1; 4;
KW   5-triphosphate 5-phosphatase type 1; IRAK1; IRAK4; KIAA1441; MSTP030;
KW   Nek9; PAR3; Pellino 1; Pellino 3; podocalyxin-like protein 1 precursor;
KW   Pushover; S-adenosylhomocysteinase;
KW   secretory carrier-associated membrane protein 2; surfeit locus protein 2;
KW   ubiquitin carboxyl terminal hydrolase 11;
KW   upstream regulatory element binding protein 1; Vartul;
KW   Werner's syndrome helicase interacting protein; WHIP;
KW   X-ray repair cross complementing protein 4.
XX
OS   Homo sapiens.
XX
PN   WO2004031242-A2.
XX
PD   15-APR-2004.
XX
PF   11-SEP-2003; 2003WO-EP010110.
XX
PR   12-SEP-2002; 2002EP-00020495.
PR   12-SEP-2002; 2002EP-00020496.
PR   12-SEP-2002; 2002EP-00020497.
XX
PA   (CELL-) CELLZOME AG.
XX
PI   Bouwmeester T,  Drewes G,  Jackson D,  Helftenbein G,  Schirle M;
PI   Kuester B,  Hopf C;
XX
DR   WPI; 2004-316467/29.
XX
PT   New complex comprising at least one first protein, and at least one
PT   second protein, useful for treating stroke, Alzheimer's disease,
PT   neurological disorders such as epilepsy, and inflammatory conditions such
PT   as ulcerative colitis.
XX
PS   Example; Page 191-193; 287pp; English.
XX
CC   The specification describes protein complexes involved in cellular
CC   processes which have been shown to be critical for the development of
CC   various forms of neurological diseases. Three protein complexes were
CC   identified: ASK2 protein complex, Pellino-1 protein complex and Pellino-3
CC   protein complex. The protein complex are useful for treating diseases and
CC   disorders, e.g. stroke, neurodegeneration such as Wallerian degeneration,
CC   Alzheimer's disease, neurological disorders such as epilepsy, and
CC   inflammatory conditions such as ulcerative colitis, Crohn's disease or
CC   atherosclerosis. Proteins identified as being part of the protein
CC   complexes of the invention are 1D-MYO-inositol triphosphate 3 kinase A,
CC   ASK1, ASK2, ASK3, CaMKII beta, CaMKII delta, CaMKII gamma, casein kinase
CC   II alpha, Cdc37, CHK2, CTCL tumour antigen SE20-4, EF-1 alpha 1, EMAP,
CC   FLJ14653 NT2RP2002252, FLJ30839 FEBRA2002429, HERC2, two hypothetical
CC   proteins of 35.5 Kda and 49.3 Kda, inositol polyphosphate-5-phosphatase,
CC   inositol-1,4,5-triphosphate 5-phosphatase type 1, IRAK1, IRAK4, KIAA1441,
CC   MSTP030, Nek9, PAR3, Pellino 1, Pellino 3, podocalyxin-like protein 1
CC   precursor, Pushover, a putative S-adenosylhomocysteinase, secretory
CC   carrier-associated membrane protein 2, surfeit locus protein 2, ubiquitin
CC   carboxyl terminal hydrolase 11, upstream regulatory element binding
CC   protein 1, Vartul, Werner's syndrome helicase interacting protein (WHIP),
CC   X-ray repair cross complementing protein 4 (isoform 1). The present
CC   sequence represents putative S-adenosylhomocysteinase.
XX
SQ   Sequence 597 AA;

  Query Match             92.8%;  Score 2582;  DB 8;  Length 597;
  Best Local Similarity   88.8%;  Pred. No. 7.7e-248;
  Matches  502;  Conservative    8;  Mismatches   19;  Indels   36;  Gaps    4;

Qy          1 MSMPDAMPLP-------GVGE---ELKQAKEIEDAEKYSFMATVTKAPKK---------- 40
              :| | |:|||       |  |    |:| :  | ::    :  :  |
Db         34 ISQP-ALPLPQETQAPSGESENLRNLRQGRRAESSQNLLSLRRLVSAQGNGRSDWRWLWI 92

Qy         41 --------------QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSS 85
                            ||||||||||||||||||||||||||||||||||||||||||||
Db         93 QLWPEIDSLEISCFLQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSS 152

Qy         86 DDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAG 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 DDEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAG 212

Qy        146 AKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESE 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 AKIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESE 272

Qy        206 DDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQL 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 DDFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQL 332

Qy        266 SKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKG 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 SKAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKG 392

Qy        326 CCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHL 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 CCAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHL 452

Qy        386 DRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNL 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 DRMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNL 512

Qy        446 SCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLT 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 SCSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLT 572

Qy        506 ELTDDQAKYLGLNKNGPFKPNYYRY 530
              |||||||||||||||||||||||||
Db        573 ELTDDQAKYLGLNKNGPFKPNYYRY 597


RESULT 11
ABM80935
ID   ABM80935 standard; protein; 611 AA.
XX
AC   ABM80935;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Tumour-associated antigenic target (TAT) polypeptide PRO81608, SEQ:2410.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   chromosome identification; chromosome mapping; gene mapping;
KW   gene therapy; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wu TD,  Zhang Z,  Zhou Y;
XX
DR   WPI; 2004-347921/32.
DR   N-PSDB; ACN38718.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
```

```
XX
PS   Claim 12; SEQ ID NO 2410; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)
CC   polypeptides, and their related nucleic acids. The TAT polypeptides are
CC   overexpressed in cancer tissues compared to normal tissues, and may thus
CC   serve as effective targets for the diagnosis and treatment of cancer in
CC   mammals. The invention also relates to nucleic acid and polypeptide
CC   sequences at least 80% identical to the TAT nucleic acids and
CC   polypeptides; expression vectors and host cells comprising a TAT nucleic
CC   acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC   molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC   TAT polypeptide; and methods and compositions for the treatment or
CC   diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC   antibodies, antagonists, binding molecules and compositions are useful
CC   for diagnosing or treating a cell proliferative disorder associated with
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT polypeptide of the invention
XX
SQ   Sequence 611 AA;

  Query Match             88.3%;  Score 2456;  DB 8;  Length 611;
  Best Local Similarity   84.9%;  Pred. No. 3e-235;
  Matches  468;  Conservative   31;  Mismatches   24;  Indels   28;  Gaps    3;

Qy          8 PLPGVGE-----------ELKQAKEIEDAEKY-------------SFMATVTKAP---K 39
              |:|| |              :: ||  :: :: :             |   |||:||   |
Db         61 PVPGPGSGPAAALSPAAGKVPQASAMKRSDPHHQHQRHRDGGEALVSPDGTVTEAPRTVK 120

Qy         40 KQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNS 99
              |||||||  ||| | ||| |||||||||||||||||||||||||||||||||||| |||:||| ||
Db        121 KQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNS 180

Qy        100 KGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTA 159
              ||||:|||||||||||||||||||:| ||::|||||||||||||||||||||||||||||
Db        181 KGSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGAKIVGCTHITAQTA 240

Qy        160 VLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMD 219
              ||:||| |||||||:|||||| |||||||||:|  |||||||||||||||||||||::
Db        241 VLMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGESEDDFWWCIDRCVNVE 300

Qy        220 GWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVN 279
              ||| ||||||||||||:|||||:||||:|||||||||||||||||||||||||||||||
Db        301 GWQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLSKAGKLCVPAMNVN 360

Qy        280 DSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYI 339
              |||||||||||||||||||||||:|||||||||||||||||||||||:|:|||:
Db        361 DSVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGCCAALKAMGSIVYV 420

Qy        340 TEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGH 399
              |||||||||||||||:|||||||||:|||||||||||||||||||||||||||||||||
Db        421 TEIDPICALQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLDRMKNSCIVCNMGH 480

Qy        400 SNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITA 459
              ||||||| |||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        481 SNTEIDVASLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLNLSCSTVPTFVLSITA 540

Qy        460 TTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNK 519
              |||||||||||||||||||||||||||||||||||:||||||||:|||||||||||||||
Db        541 TTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTELTDEQAKYLGLNK 600

Qy        520 NGPFKPNYYRY 530
              ||||||||||
Db        601 NGPFKPNYYRY 611


RESULT 12
ADS10638
ID   ADS10638 standard; protein; 508 AA.
XX
AC   ADS10638;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Human therapeutic protein - SEQ ID 875.
XX
KW   antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
KW   inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
KW   aplastic anaemia; cancer; wound healing; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2004080148-A2.
XX
PD   23-SEP-2004.
XX
PF   30-SEP-2003; 2003WO-US030720.
XX
PR   02-OCT-2002; 2002US-0416186P.
XX
PA   (NUVE-) NUVELO INC.
XX
PI   Tang YT,  Asundi V,  Ren F,  Zhang J,  Wehrman T,  Wang Z,  Ma Y;
PI   Wang D,  Chen R,  Zhao QA,  Wang J,  Ghosh M,  Xue AJ,  Weng G,  Zhou P;
XX
DR   WPI; 2004-668857/65.
DR   N-PSDB; ADS09954.
XX
PT   New polynucleotide, useful in preparing a composition for diagnosing or
PT   treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT   aplastic anemia or cancer for promoting wound healing.
XX
PS   Claim 20; SEQ ID NO 875; 718pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide and the encoded
CC   polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC   neuroprotective, antianaemic, cytostatic and vulnerary activities and may
CC   be useful in preparing a composition for diagnosing or treating
CC   inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC   disorders, such as aplastic anaemia or cancer, as well as for promoting
CC   wound healing. The molecules may also be utilised during gene therapy
CC   procedures. The current sequence is that of a human therapeutic protein
CC   of the invention. The current sequence is not shown explicitly within the
CC   specification but can be accessed from the WIPO web-site.
XX
SQ   Sequence 508 AA;

  Query Match             87.3%;  Score 2429;  DB 8;  Length 508;
  Best Local Similarity   92.1%;  Pred. No. 1.1e-232;
  Matches  454;  Conservative   23;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         38 PKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQT 97
              |:  |||||  ||| | ||| |||||||||||||||||||||||||||||||||||:|||
Db         16 PEWMIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQK 75

Qy         98 NSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQ 157
              ||||||:|||||||||||||||||||:| ||::|||||||||||||||||||||||||||
Db         76 NSKGSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGAKIVGCTHITAQ 135

Qy        158 TAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVN 217
              ||||:||| |||||||:|||||| |||||||||:|  |||||||||||||||||||||
Db        136 TAVLMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGESEDDFWWCIDRCVN 195

Qy        218 MDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMN 277
              ::||| |||||||||||:||||||:||||:||||||||||||||||||||||||||||||
Db        196 VEGWQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLSKAGKLCVPAMN 255

Qy        278 VNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIV 337
              |||||||||||||||||||||||||:|||||||||||||||||||||||:|:||
Db        256 VNDSVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGCCAALKAMGSIV 315
```

```
Qy        338 YITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNM 397
              |:|||||||||||||||||:||||||||||:|||||||||||||||||||||||||||||
Db        316 YVTEIDPICALQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLDRMKNSCIVCNM 375

Qy        398 GHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSI 457
              |||||||||| ||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        376 GHSNTEIDVASLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLNLSCSTVPTFVLSI 435

Qy        458 TATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGL 517
              |||||||||||||||||||||||||||||||||||||||:||||||||||:|||||||
Db        436 TATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTELTDEQAKYLGL 495

Qy        518 NKNGPFKPNYYRY 530
              |||||||||||||
Db        496 NKNGPFKPNYYRY 508


RESULT 13
ADL90193
ID   ADL90193 standard; protein; 509 AA.
XX
AC   ADL90193;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Human enzyme, ENZM-3, SEQ ID 3.
XX
KW   Cytostatic; Antiarteriosclerotic; Anticonvulsant; Nootropic;
KW   Neuroprotective; Cerebroprotective; Anti-HIV; Antiallergic;
KW   Antiinflammatory; Thyromimetic; Gene therapy; human; enzyme;
KW   cell proliferative disorder; cancer; atherosclerosis;
KW   neurological disorder; epilepsy; Huntington's disease; stroke;
KW   immune disorder; inflammatory disorder; AIDS; allergies;
KW   developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
KW   ENZM-3.
XX
OS   Homo sapiens.
XX
PN   WO2004027022-A2.
XX
PD   01-APR-2004.
XX
PF   05-SEP-2003; 2003WO-US028177.
XX
PR   05-SEP-2002; 2002US-0408747P.
PR   28-OCT-2002; 2002US-0422062P.
PR   29-OCT-2002; 2002US-0422276P.
PR   13-DEC-2002; 2002US-0433328P.
PR   13-JAN-2003; 2003US-0439997P.
PR   15-JAN-2003; 2003US-0440850P.
PR   31-JAN-2003; 2003US-0444282P.
PR   04-FEB-2003; 2003US-0445371P.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Sanjanwala MM,  Lee S,  Lee SY,  Tran UK,  Lu Y,  Baughn MR;
PI   Chawla NK,  Lal PG,  Ring HZ,  Yang YG,  Hafalia AJA,  Yao MG;
PI   Swarnakar A,  Ison CH,  Chang H,  Ramkumar J,  Khare R,  Bhatia UG;
PI   Burrill JD,  Blake JJ,  Ho A,  Zheng W,  Jiang X,  Jackson AA;
PI   Marquis JP,  Jin P,  Wilson AD,  Favero KD,  Wang JT,  Becha SD;
PI   Naidu S,  Yue H,  Griffin JA,  Kable AE,  Emerling BM,  Lee EA,  Tang YT;
PI   Li JX,  Forsythe IJ;
XX
DR   WPI; 2004-295399/27.
DR   N-PSDB; ADL90239.
XX
PT   New human enzymes (ENZM), useful for diagnosing, treating and preventing
PT   diseases or conditions associated with the aberrant ENZM expression e.g.
PT   cancer, AIDS, epilepsy, or infections.
XX
PS   Claim 1; SEQ ID NO 3; 344pp; English.
XX
CC   The present invention relates to human enzymes (ENZM-1 - ENZM-46,
CC   ADL90191-ADL90236) and their coding sequences (ADL90237-ADL90282). The
CC   sequences are useful in diagnosing, treating and preventing diseases or
CC   conditions associated with the decreased expression or overexpression of
CC   the enzymes, such as cell proliferative (e.g. cancer, atherosclerosis),
CC   neurological (e.g. epilepsy, Huntington's disease, stroke),
CC   immune/inflammatory (e.g. AIDS, allergies) and developmental (e.g.
CC   Hypothyroidism, Cushing's syndrome) disorders, or infections.
XX
SQ   Sequence 509 AA;

  Query Match             87.3%;  Score 2429;  DB 8;  Length 509;
  Best Local Similarity   92.7%;  Pred. No. 1.1e-232;
  Matches  454;  Conservative   22;  Mismatches   14;  Indels    0;  Gaps    0;

Qy         41 QIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSDDEVSPREKQQTNSK 100
              |||||| ||| | ||| |||||||||||||||||||||||||||||| |||:||| |||
Db         20 QIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSDDETSPRDKQQKNSK 79

Qy        101 GSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAV 160
              |||:|||||||||||||||||||:| ||::|||||||||||||||||||||||||||||
Db         80 GSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGAKIVGCTHITAQTAV 139

Qy        161 LIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDG 220
              |:||| |||||||:|||||| |||||||||:|  ||||||||||||||||||||::|
Db        140 LMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGESEDDFWWCIDRCVNVEG 199

Qy        221 WQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVND 280
              || |||||||||||:|||||:||||:|||||||||||||||||||||||||||||||||
Db        200 WQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLSKAGKLCVPAMNVND 259

Qy        281 SVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYIT 340
              |||||||||||||||||||||||:|||||||||||||||||||||||||:|:|||:|
Db        260 SVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGCCAALKAMGSIVYVT 319

Qy        341 EIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHS 400
              |||||||||||||||:||||||||||:||||||||||||||||||||||||||||||||
Db        320 EIDPICALQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLDRMKNSCIVCNMGHS 379

Qy        401 NTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITAT 460
              |||||| ||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        380 NTEIDVASLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLNLSCSTVPTFVLSITAT 439

Qy        461 TQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKN 520
              |||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||
Db        440 TQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTELTDEQAKYLGLNKN 499

Qy        521 GPFKPNYYRY 530
              ||||||||||
Db        500 GPFKPNYYRY 509


RESULT 14
ABG32371
ID   ABG32371 standard; protein; 564 AA.
XX
AC   ABG32371;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Novel human enzyme, NZMS-1, Incyte 7484737CD1.
XX
KW   Human; NZMS-1; immune system disorder; AIDS; contact dermatitis;
KW   acquired immune deficiency syndrome; Chediak-Higashi syndrome;
KW   common variable immunodeficiency; developmental disorder; cataract;
KW   Cushing's syndrome; hypothyroidism; eye disorder; glaucoma; asthma;
KW   metabolic disorder; cystic fibrosis; sickle-cell anaemia; epilepsy;
KW   smooth muscle disorder; lactic acidosis; neurological disorder;
KW   Parkinson's disease; cardiac disorder; congestive heart failure;
KW   myocardial infarction; pulmonary disorder; emphysema; bronchitis;
KW   parasitic infection; amoebiasis; filariasis; cancer; leukaemia;
KW   cell proliferative disorder; Incyte 7484737CD1; enzyme.
XX
OS   Homo sapiens.
```

XX
PN WO200264795-A2.
XX
PD 22-AUG-2002.
XX
PF 08-FEB-2002; 2002WO-US003814.
XX
PR 09-FEB-2001; 2001US-0268113P.
PR 15-FEB-2001; 2001US-0269215P.
PR 27-FEB-2001; 2001US-0272271P.
PR 07-MAR-2001; 2001US-0274091P.
PR 09-MAR-2001; 2001US-0274423P.
PR 23-MAR-2001; 2001US-0278479P.
PR 23-MAR-2001; 2001US-0278480P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Sanjanwala MM, Lu Y, Lee EA, Hafalia AJA, Warren BA, Baughn MR;
PI Tang TY, Yue H, Yao MG, Lee S, Thornton M, Walia NK, Xu Y;
PI Tran UK, Lal PG, Lu DAM, Swarnaker A, Jones KA, Ring HZ;
XX
DR WPI; 2002-643488/69.
DR N-PSDB; ABK91312.
XX
PT New isolated human enzymes (NZMS), useful for diagnosing, treating and
PT preventing immune system, developmental, eye, metabolic, smooth muscle,
PT neurological, pulmonary and cardiac disorders, cancer, or parasitic
PT infections.
XX
PS Claim 1; Page 147-149; 170pp; English.
XX
CC The invention relates to a new isolated human enzyme (NZMS) (I) and the
CC encoding nucleic acid (II). The polypeptides and polynucleotides are
CC useful in assessing the effects of exogenous compounds on the expression
CC of nucleic acid and amino acid sequences of NZMS. The NZMS or its
CC fragments are useful in screening compounds for effectiveness as agonist
CC or antagonist of the polypeptides, or in altering the expression of the
CC target polynucleotide and compounds that specifically bind to or modulate
CC the activity of the polypeptide. The antibody is useful in diagnosing a
CC condition or disease associated with the decreased expression or
CC overexpression of NZMS and in detecting (I). (I) is useful in the
CC production of antibodies. The microarray is useful in monitoring or
CC measuring gene expression profiles. The polypeptides and polynucleotides
CC are useful in diagnosing, treating and preventing diseases or conditions
CC associated with the decreased expression or overexpression of NZMS, such
CC as immune system disorders (e.g. acquired immune deficiency syndrome
CC (AIDS), contact dermatitis), immune deficiencies (e.g. common variable
CC immunodeficiency, Chediak-Higashi syndrome), developmental disorders
CC (e.g. Cushing's syndrome, hypothyroidism), eye disorders (e.g. glaucoma,
CC cataract), metabolic disorders (e.g. cystic fibrosis, sickle-cell
CC anaemia), smooth muscle disorders (e.g. asthma, lactic acidosis),
CC neurological disorders (e.g. Parkinson's disease, epilepsy), cardiac
CC disorders (e.g. congestive heart failure, myocardial infarction),
CC pulmonary disorders (emphysema, bronchitis), parasitic infections
CC (amoebiasis, filariasis), and cell proliferative disorders (e.g. cancer,
CC leukaemia). The microarray is useful in monitoring or measuring protein-
CC protein interactions, drug-target interactions, and gene expression
CC profiles. The present sequence represents the amino acid sequence of one
CC of 11 novel human enzymes (NZMS1-11) of the invention
XX
SQ Sequence 564 AA;

Query Match 87.3%; Score 2429; DB 5; Length 564;
Best Local Similarity 90.3%; Pred. No. 1.3e-232;
Matches 455; Conservative 24; Mismatches 25; Indels 0; Gaps 0;

```
Qy         27 KYSFMATVTKAPKKQIQFADDMQEFTKFPTKTGRRSLSRSISQSSTDSYSSAASYTDSSD 86
              |  ::      | ||||||  ||| | ||| ||||||||||||||||||||||||||||
Db         61 KKKYIVNGNSGIKAQIQFADQKQEFNKRPTKIGRRSLSRSISQSSTDSYSSAASYTDSSD 120

Qy         87 DEVSPREKQQTNSKGSSNFCVKNIKQAEFGRREIEIAEQDMSALISLRKRAQGEKPLAGA 146
              || |||:||| |||||||:|||||||||||||||||||||:| ||::|||||||||||||
Db        121 DETSPRDKQQKNSKGSSDFCVKNIKQAEFGRREIEIAEQEMPALMALRKRAQGEKPLAGA 180

Qy        147 KIVGCTHITAQTAVLIETLCALGAQCRWSACNIYSTQNEVAAALAEAGVAVFAWKGESED 206
              |||||||||||||||:||| |||||||:||||||| |||||||||:|  ||||||||||
Db        181 KIVGCTHITAQTAVLMETLGALGAQCRWAACNIYSTLNEVAAALAESGFPVFAWKGESED 240

Qy        207 DFWWCIDRCVNMDGWQANMILDDGGDLTHWVYKKYPNVFKKIRGIVEESVTGVHRLYQLS 266
              ||||||||||::||| |||||||||||||:||||||:||||:||||||||||||||||||
Db        241 DFWWCIDRCVNVEGWQPNMILDDGGDLTHWIYKKYPNMFKKIKGIVEESVTGVHRLYQLS 300

Qy        267 KAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDVMFGGKQVVVCGYGEVGKGC 326
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        301 KAGKLCVPAMNVNDSVTKQKFDNLYCCRESILDGLKRTTDMMFGGKQVVVCGYGEVGKGC 360

Qy        327 CAALKALGAIVYITEIDPICALQACMDGFRVVKLNEVIRQVDVVITCTGNKNVVTREHLD 386
              ||||||:|:|||:|||||||||||||||:|||||||||:|||||||||||||||||||||
Db        361 CAALKAMGSIVYVTEIDPICALQACMDGFRLVKLNEVIRQVDIVITCTGNKNVVTREHLD 420

Qy        387 RMKNSCIVCNMGHSNTEIDVTSLRTPELTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLS 446
              |||||||||||||||| || |||||||||||||||||||||||:|||||||||||||||
Db        421 RMKNSCIVCNMGHSNTETDVASLRTPELTWERVRSQVDHVIWPDGKRIVLLAEGRLLNLS 480

Qy        447 CSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTE 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db        481 CSTVPTFVLSITATTQALALIELYNAPEGRYKQDVYLLPKKMDEYVASLHLPTFDAHLTE 540

Qy        507 LTDDQAKYLGLNKNGPFKPNYYRY 530
              |||:||||||||||||||||||||
Db        541 LTDEQAKYLGLNKNGPFKPNYYRY 564
```

RESULT 15
ABB58663
ID ABB58663 standard; protein; 521 AA.
XX
AC ABB58663;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 2781.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL02766.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 2781; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 521 AA;

Query Match 70.8%; Score 1970; DB 4; Length 521;
Best Local Similarity 77.6%; Pred. No. 7.4e-187;
Matches 370; Conservative 48; Mismatches 55; Indels 4; Gaps 3;

```
Qy         57 KTGRRSLSRSISQSSTDSYSSAASYTDSSD--DEVSPREKQQTNSKGSSNFCVKNI-KQA 113
              |   |  |||:| |||||:|| |||| ||:  |:| |||| | ||||||:|||:||  |
Db         46 KKTSRYRSRSLSASSTDSFSS-ASYTGSSEDGDDVPPREKVQKNSKGSSDFCVRNIGAQH 104

Qy        114 EFGRREIEIAEQDMSALISLRKRAQGEKPLAGAKIVGCTHITAQTAVLIETLCALGAQCR 173
              ||||||||||:|  :|:|:|||  :|||  ||||||||| |||||||||  |||  |
Db        105 AFGRREIEIAEQEMPGIIALKKRAAEDKPLKDAKIVGCTHINAQTAVLIETLVELGASVR 164

Qy        174 WSACNIYSTQNEVAAALAEAGVAVFAWKGESEDDFWWCIDRCVNMDGWQANMILDDGGDL 233
              |:|||||||||||||||:|: :|||:||:|:|||||||||| : || |||||||||
Db        165 WAACNIYSTQNEVAAALAESGIPIFAWRGETEEDFWWCIDRCVNAENWQPNMILDDGGDA 224

Qy        234 THWVYKKYPNVFKKIRGIVEESVTGVHRLYQLSKAGKLCVPAMNVNDSVTKQKFDNLYCC 293
              || : |||| :|| ::|||||||||||||||||||||| |||||||||||| |||||| |
Db        225 THLMLKKYPTMFKLVKGIVEESVTGVHRLYQLSKAGKLTVPAMNVNDSVTKTKFDNLYSC 284

Qy        294 RESILDGLKRTTDVMFGGKQVVVCGYGEVGKGCCAALKALGAIVYITEIDPICALQACMD 353
              :||||| |||:|||||||||||||||:|||||  |||  | |||||||||||||| ||
Db        285 KESILDSLKRSTDVMFGGKQVVVCGYGDVGKGCAQALKGQGCIVYITEIDPICALQASMD 344

Qy        354 GFRVVKLNEVIRQVDVVITCTGNKNVVTREHLDRMKNSCIVCNMGHSNTEIDVTSLRTPE 413
              |||||||||||| ||:|:| ||||||| |||:|:||: |||||||||||||  ||||:
Db        345 GFRVVKLNEVIRNVDIVVTATGNKNVVVREHMDKMKSGCIVCNMGHSNTEIDVNGLRTPD 404

Qy        414 LTWERVRSQVDHVIWPDGKRVVLLAEGRLLNLSCSTVPTFVLSITATTQALALIELYNAP 473
              ||||:|||||||:|||:|| ::|||||||:||||||::|:| :|||: |||||||||:|||
Db        405 LTWEKVRSQVDHIIWPEGKYIILLAEGRLVNLSCSSIPSFAVSITSATQALALIELFNAP 464

Qy        474 EGRYKQDVYLLPKKMDEYVASLHLPSFDAHLTELTDDQAKYLGLNKNGPFKPNYYRY 530
               |||| |||||||||||||||:|||||||:|:|||||:|||| ||||||||||
Db        465 PGRYKSDVYLLPKKMDEYVASLHLPTFDAHLTELSDEQAKYMGLNKAGPFKPNYYRY 521
```

Search completed: December 14, 2005, 08:16:40
Job time : 135.918 secs